

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:36:05 ; Search time 610.591 Seconds
(without alignments)
756.292 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103
Sequence: 1 DPEFTPEPVPNSIGKSPDS 19

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -dbv=xip
-Q/cgn2_1/USPTO.spool/US100000039/runat_01082003_085419_27521/app_query.fasta_1.661
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.csl -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US100000039 -CGN_1_1_5635 -runat_01082003_085419_27521 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estlin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_psq.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	259	9	AA344258
2	103	100.0	287	14	N58770
3	103	100.0	344	10	BE001610
4	103	100.0	346	9	AA377138
5	103	100.0	376	12	B1032559
6	103	100.0	404	14	H72913
7	103	100.0	420	14	R86234
8	103	100.0	424	14	N49311
9	103	100.0	427	10	BE814167
10	103	100.0	450	14	H51048
11	103	100.0	460	14	R35214
12	103	100.0	465	10	BE002651
13	103	100.0	465	12	BM704999
14	103	100.0	505	12	BM707970
15	103	100.0	530	12	BM756001
16	103	100.0	544	13	BQ083706
17	103	100.0	557	12	BM705333
18	103	100.0	561	14	CB154488
19	103	100.0	572	12	BM759618
20	103	100.0	586	12	B1003243
21	103	100.0	590	12	BM842453
22	103	100.0	608	14	CB216184
23	103	100.0	620	10	BF350342
24	103	100.0	630	14	CB216592
25	103	100.0	641	9	AW378705
26	103	100.0	654	9	AL040857
27	103	100.0	659	10	BE178455
28	103	100.0	661	10	BE178442
29	103	100.0	663	10	BE178469
30	103	100.0	671	12	BE770190
31	103	100.0	703	10	BG485236
32	103	100.0	721	12	BG778994
33	103	100.0	722	9	AV703365
34	103	100.0	739	14	CB958767
35	103	100.0	765	12	BG763506
36	103	100.0	767	14	CD000714
37	103	100.0	826	14	CD517205
38	103	100.0	834	12	B1335545
39	103	100.0	847	13	BQ212589
40	103	100.0	851	12	B1764996
41	103	100.0	859	13	BQ680126
42	103	100.0	873	10	BG756781
43	103	100.0	876	10	BG122546
44	103	100.0	898	13	BQ687393
45	103	100.0	899	12	BG769533

ALIGNMENTS

RESULT 1
AA344258

LOCUS
DEFINITION

AA344258

Accession

AA344258.1

Keywords

EST.

Source

Homo sapiens

Organism

Homo sapiens

Reference

1 (bases 1 to 259)

259 bp mRNA linear EST 21-APR-1997

EST50139

Gall bladder 1 Homo sapiens CDNA 5' end similar to similar

to serine/threonine kinase, mRNA sequence.

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
C.J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Cocayne, J.D., White
O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wal, C., Clayton, R.A.,
Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald,
L.M., Fitzhugh, W.M., Fritchman, J.L., Geochagen, N.S., Glodok, A.,
Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
Kelley, J.C., Liu, L.-I., Marmaro, S.M., Metrick, J.M.,
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,
Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,
Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
Dinke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.,
Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,
Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, R., Raymond, L.,
Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon
M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
Venter, J.C.

TITLE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

JOURNAL
Nature 377 (6547 Suppl), 3-174 (1995)

MEDLINE
96026280

PUBMED
7566098

COMMENT
Other ESTs: THC173211

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hqi/hqi.html>)

Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers

source

1..259

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (Inhost):145789"

/db_xref="taxon:9606"

/sex="female"

/dev_stage="adult, 25 yrs"

/clone_lib="Gall bladder 1"

/note="Organ: gall bladder; Vector: pBluescript SK-"

Site_1: EcoRI; Site_2: XhoI

BASE COUNT 49 a 77 c 60 g 72 t

ORIGIN

Alignment Scores:

Prod. No.:	5,71e-06	Length:	259
Score:	103.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-000-039a-3 (1-19) x AA344258 (1-259)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

DB 49 GACCCCGAGTTTACCGAGAGCGCTGTCCTCCCACTCCATTCGACAGTCCCTGACAGC 105

RESULT 2

N58770

LOCUS N58770 287 bp mRNA linear EST 28-JAN-1997

DEFINITION yv75f11.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone

IMAGE:248589 3' similar to SW:SGK_RAY Q06226

SERINE/THREONINE-PROTEIN KINASE SGK 1, mRNA sequence.

ACCESSION N58770

VERSION N58770.1 GI:1202660

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 287)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubucque, T., Favellio, A., Gish, W., Hawkins,
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore,
B., Morris, M., Parsons, J., Prange, C., Rifkin, B., Kohlman, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E.,
Underwood, K., Wohlmann, P., Waterston, K., Wilson, R. and Maria, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Wilson KK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through BLIM; contact the

IMAGE Consortium (image.lim.gov) for further information.

Possible reversed clone: similarity on wrong strand

Insert Length: 1014 Std Error: 0.00

Seq primer: mob.RPGA+Er

High quality sequence stop: 1.

FEATURES
Location/Qualifiers

1..287

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:3797835"

/db_xref="taxon:9606"

/clone="IMAGE:248589"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/clone_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: p774p (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5'-AACTGGAAGATTAATAAGATCTTTTTTTTTTTTTTTT 3']

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified p773 vector. Library

went through one round of normalization. Library

constructed by Beato Soares and M.Fatima Khalid.

BASE COUNT 59 a 84 c 63 g 79 t

ORIGIN

Alignment Scores:

Prod. No.:	6,42e-06	Length:	287
Score:	103.00	Matches:	19
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-000-039a-3 (1-19) x N58770 (1-287)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

DB 70 GACCCCGAGTTTACCGAGAGCGCTGTCCTCCCACTCCATTCGACAGTCCCTGACAGC 126

RESULT 3

BE001610

LOCUS BE001610 344 bp mRNA linear EST 05-JUN-2000

DEFINITION FM2-BN0080-270400-003-a12 BN0080 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE001610

VERSION BE001610.1 GI:8261843

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE:
AUTHORS
1 (bases 1 to 344)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Priones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
CONTACT: Simpson A.J.C.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-PW2-BN0080-270
400-003-a12at3-2000-04-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 30
High quality sequence stop: 344.
Location/Qualifiers
1. 344
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="BN0080"
/note="Organ: breast normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 75 a 92 c 76 g 101 t
ORIGIN
Alignment Scores:
Pred. No.: 7.9e-06 Length: 344
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-10-000-039a-3 (1-19) x BE001610 (1-344)
QY 1 AspGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 137 GACCCCGAGTTTACCGAGAGCCCTGTCCCAACTCCATTCGCAAGTCCCTGCAGC 193
|||||

RESULT 4
AA377138
LOCUS
DEFINITION
EST89672 Small intestine I Homo sapiens cDNA 5' end similar to similar to serine/threonine kinase, mRNA sequence.
ACCESSION
AA377138
VERSION
AA377138.1 GI:2029466
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 346)
Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Feldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White

O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.B., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodzik, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L., I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelletier, S.M., Phillips, C.A., Ryder, S.F., Scott, J.L., Saudek, D.M., Shetty, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, F.A., Collins, E.J., Dinko, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, I., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseitine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
6026280
7566098
Other_ESTs: THC173211
Contact: Kerlavage, A.R.
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavage@igrr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (http://www.tigr.org/tldb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 346
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):181556"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="Small intestine I"
/note="Organ: small intestine; Vector: pBluescript SK-
Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 80 a 93 c 74 g 97 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 7.95e-06 Length: 346
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-000-039a-3 (1-19) x AA377138 (1-346)
QY 1 AspGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 67 GACCCCGAGTTTACCGAGAGCCCTGTCCCAACTCCATTCGCAAGTCCCTGCAGC 123
|||||

RESULT 5
BI032559
LOCUS
DEFINITION
CM3-NN0246-120101-535-b07 NN0246 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BI032559
VERSION
BI032559.1 GI:14439185
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 376)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Priones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P., and Wilson, R.
The WashU-Merck EST Project
Unpublished

TITLE
JOURNAL
COMMENT

Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 1276

High quality sequence stops: 128

Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Insert length: 1276 Std Error: 0.00

Seq primer: Promega -21ml3

High quality sequence stop: 128.

Location/Qualifiers

1. .420

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:3763564"

/db_xref="taxon:9606"

/clone="IMAGE:194514"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="PH108 (ampicillin resistant)"

/clone_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - Oligo(dT) primer
[5' AACCGGAGAAATTAATTAAGATCTTTTTTTTTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 87 a 111 c 102 g 118 t 2 others

ORIGIN

Alignment Scores:
Pred. No.: 9,91e-06 Length: 420
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-000-039A-3 (1-19) x R86234 (1-420)

QY 1 AspproGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

|||||
DB 39 GACCCCGAGTTTACCGAAGAGCTGTCCCACTCCATTGGCAAGTCCCTGACACG 95

RESULT 8

N49311

LOCUS

DEFINITION

YV24c11.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone

IMAGE:243668 3', similar to SW:SGK_RAT Q06226

SERINE/THREONINE-PROTEIN KINASE SGK ; mRNA sequence.

N49311 GI:1190477

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

EST.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 424)

REFERENCE

AUTHORS

Hilf, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,

M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,

R., Williamson, A., Wohldmann, P., and Wilson, R.

TITLE

JOURNAL

COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Possible reversed clone: similarity on wrong strand

Seq primer: m13 -40 forward

High quality sequence stop: 402.

Location/Qualifiers

1. .424

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:4792801"

/db_xref="taxon:9606"

/clone="IMAGE:243668"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="PH108 (ampicillin resistant)"

/clone_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: p7714b (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5' AACCGGAGAAATTAATTAAGATCTTTTTTTTTTTT 3']

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified p773 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 92 a 114 c 89 g 129 t

ORIGIN

Alignment Scores:

Pred. No.: 1e-05 Length: 424

Score: 103.00 Matches: 19

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-10-000-039A-3 (1-19) x N49311 (1-424)

QY 1 AspproGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

|||||
DB 71 GACCCCGAGTTTACCGAAGAGCTGTCCCACTCCATTGGCAAGTCCCTGACACG 127

RESULT 9

BE814167

LOCUS

DEFINITION

PMO-BN0066-130600-004-b02 BN0066 Homo sapiens cDNA, mRNA sequence.

ACCESSION

BE814167

VERSION

BE814167.1

KEYWORDS

EST.

Source

Homo sapiens (human)

EST.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 427)

REFERENCE

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,

Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, P. F.,

Goldman, G. H., Carvalho, A. F., Matsukuma, A., Bata, G. S., Simpson, D. H.,

Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare

M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and

Simpson, A. J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)


```

source
1. .465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agg-9-13-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="PH108 (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/notes="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I and the (dT)18 tail. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
```

BASE COUNT 101 a 124 c 101 g 139 t

Alignment Scores: 1.11e-05 Length: 465

Pred. No.: 103.00 Matches: 19

Score: 100.00 Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00% Indels: 0

Query Match: 100.00% Gaps: 0

DB: 12

US-10-000-039a-3 (1-19) x BM704999 (1-465)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

Db 67 GACCCGAGTTTACCGAGAGCTGTCCCACTCATTTGGCAAGTCCCTGACAGC 123

RESULT 14

BM707970

LOCUS

DEFINITION

UI-E-C11-akt-k17-0-UI.r1 UI-E-C11 Homo sapiens cDNA clone

BM707970

VERSION

KEYWORDS

SOURCE

EST.

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Bonaldo,M.F., Lennon,G. and Soares,M.B.

1 (bases 1 to 505)

Normalization and subtraction: two approaches to facilitate gene discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

CONTACT: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1. .505

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-C11-akt-k17-0-UI"

/tissue_type="RPE and Choroid"

/dev_stage="adult"

/lab_host="PH108 (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-C11"

/notes="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 124 a 143 c 117 g 121 t

Alignment Scores: 1.22e-05 Length: 505

Pred. No.: 103.00 Matches: 19

Score: 100.00 Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00% Indels: 0

Query Match: 100.00% Gaps: 0

DB: 12

US-10-000-039a-3 (1-19) x BM707970 (1-505)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19

Db 365 GACCCGAGTTTACCGAGAGCTGTCCCACTCATTTGGCAAGTCCCTGACAGC 421

RESULT 15

BM756001

LOCUS

DEFINITION

K-EST0034163 S1SN05 Homo sapiens cDNA clone S1SN05-38-Fin 5', mRNA

BM756001

VERSION

KEYWORDS

SOURCE

EST.

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 38 row: E column: 04

High quality sequence stop: 530.

Location/Qualifiers

1. .530

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNU5-38-E04"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="top10F"
/clone_lib="S1SNU5"
/notes="Organ: Stomach; vector: pcNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector, the dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT      125 a   140 c   120 g   145 t
ORIGIN

```

```

Alignment Scores:
Pred. No.:      1.29e-05      Length:      530
Score:          103.00      Matches:      19
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              12          Gaps:      0

US-10-000-039A-3 (1-19) x BM756001 (1-530)

QY      1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
DB      200 GACCCGAGCTTTACCGAGAGCGCTGTCCCACTCCATTGGCAGTCCCTGACAGC 256

```

Search completed: August 11, 2003, 11:27:47
Job time : 614.591 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:11:03 ; Search time 42.5451 Seconds
(without alignments)
476.401 Million cell updates/sec

Title: US-10-000-039A-2
Perfect score: 2270
Sequence: 1 MTVKTEAAKGLTYSRMRGM.....KEAAEFLGFSYAPPTDSFL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2270	100.0	431	1 SGK1_HUMAN	O00141 homo sapien
2	2215	97.6	431	1 SGK1_RABIT	Q9xt18 oryctolagus
3	2205	97.1	431	1 SGK1_MOUSE	Q9yvc6 mus musculus
4	2204.5	97.1	430	1 SGK1_RAT	Q06226 rattus norv
5	1480	65.2	496	1 SGK3_MOUSE	Q9ere3 mus musculus
6	1472	64.8	496	1 SGK3_HUMAN	Q96br1 homo sapien
7	1415	62.3	427	1 SGK2_HUMAN	Q9hby8 homo sapien
8	1383	60.9	367	1 SGK2_MOUSE	Q9qzs5 mus musculus
9	1232	54.3	302	1 SGK2_RAT	Q8r4u9 rattus norv
10	909	40.0	479	1 AKT3_HUMAN	Q9y243 homo sapien
11	909	40.0	479	1 AKT3_MOUSE	Q9wua6 mus musculus
12	903	39.8	454	1 AKT3_RAT	Q63484 rattus norv
13	865	38.1	501	1 KAKT_MLVAT	P31748 akt8 murine
14	863	38.0	480	1 KRAC_MOUSE	P31750 mus musculus
15	863	38.0	480	1 KRAC_RAT	P47196 rattus norv
16	860.5	37.9	479	1 PK2_DICDI	P28178 dictyosteli
17	857.5	37.8	444	1 KRAC_DICDI	P54644 dictyosteli
18	850	37.4	480	1 KRAC_HUMAN	P31749 homo sapien
19	849	37.4	481	1 AKT2_HUMAN	P31751 homo sapien
20	847	37.3	480	1 KRAC_BOVIN	Q01314 bos taurus
21	847	37.3	481	1 AKT2_MOUSE	P60823 mus musculus
22	838	36.9	481	1 AKT2_RAT	P47197 rattus norv
23	823.5	36.3	680	1 YPK1_YEAST	P12688 saccharomyc
24	818.5	36.1	677	1 YPK2_YEAST	P18961 saccharomyc
25	812	35.8	485	1 K6B2_MOUSE	Q9zlm4 mus musculus
26	806.5	35.5	482	1 K6B2_HUMAN	Q9ubso h ribosomal
27	803	35.4	502	1 K6B1_HUMAN	P23443 homo sapien
28	803	35.4	502	1 K6B1_RAT	P21425 rattus norv
29	795	35.0	634	1 KPC3_DROME	P13678 drosophila
30	792.5	34.9	1016	1 KPC2_SCHPO	P36583 schizosacch
31	791	34.8	672	1 KPCA_HUMAN	P17252 homo sapien
32	791	34.8	672	1 KPCA_RAT	P05696 rattus norv
33	790	34.8	672	1 KPCA_RABIT	P10102 oryctolagus

RESULT 1

ID	SGK1_HUMAN	STANDARD;	PRT:	431 AA.
AC	O00141: Q9UN56;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-Sep-2003 (Rel. 42, last annotation update)			
DE	Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)			
DE	(Serum/glucocorticoid-regulated kinase 1).			
GN	SGK OR SGK1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RX	MEDLINE=97272242; PubMed=9114008;			
RA	Waldegger S., Barth P., Raber G., Lang F.;			
RT	"Cloning and characterization of a putative human serine/threonine			
RT	protein kinase transcriptionally modified during anisotonic and			
RT	isotonic alterations of cell volume."			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:4440-4445(1997).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=98390195; PubMed=9722955;			
RA	Waldegger S., Erdel M., Nagl U.O., Barth P., Raber G., Steuer S.,			
RA	Utermann G., Paulmichl M., Lang F.;			
RT	"Genomic organization and chromosomal localization of the human SGK			
RT	protein kinase gene."			
RN	[3]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Dermal papilla;			
RC	Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Chung H.J., Sohn M.Y.,			
RA	Hwang S.Y., Im S.U., Jung E.J., Lee J.H., Kim J.C.;			
RT	"A catalogue of genes in the human dermal papilla cells as identified			
RT	by expressed sequence tags."			
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.			
RN	[4]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Cervix;			
RC	MEDLINE=22388257; PubMed=12477932;			
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.C.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Hopkins S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Altschul S.R., Jordan D., Moore T., Max J., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Abramson R.D., Mullighy S.J.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson N.C.,			

ALIGNMENTS

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length
 RA human and mouse cDNA sequences";
 RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [5]
 RL TISSUE SPECIFICITY.
 RP MEDLINE-20018032; PubMed-10548550;
 RP Kobayashi T., Deak M., Morrice N., Cohen P.;
 RA "Characterization of the structure and regulation of two novel
 RA isoforms of serum- and glucocorticoid-induced protein kinase";
 RA Biochem. J. 344:189-197(1999).
 RL [6]
 RL PHOSPHORYLATION ON THR-256, AND MUTAGENESIS OF THR-256 AND SER-422.
 RP TISSUE-BRAIN;
 RP MEDLINE-99208518; PubMed-10191262;
 RA Kobayashi T., Cohen P.;
 RA "Activation of serum- and glucocorticoid-regulated protein kinase by
 RA agonists that activate phosphatidylinositol 3-kinase is mediated by
 RA 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PK2";
 RA Biochem. J. 339:319-328(1999).
 RL [7]
 RL CHARACTERIZATION.
 RP MEDLINE-20345128; PubMed-1088438;
 RA Lang F., Klingel K., Wagner C.A., Stegen C., Waerntges S.,
 RA Friedrich B., Lanzendoerfer M., Melzig J., Moschen I., Steuer S.,
 RA Waldegger S., Sauter M., Paulmichl M., Gerke V., Riser T., Gamba G.,
 RA Capasso G., Kandolf R., Hebert S.C., Messy S.G., Broer S.;
 RA "Deranged transcriptional regulation of cell-volume-sensitive kinase
 RA hSGK in diabetic nephropathy";
 RA Proc. Natl. Acad. Sci. U.S.A. 97:8157-8162(2000).
 RL [8]
 RL FUNCTION.
 RP MEDLINE-20584913; PubMed-11154281;
 RA Brunet A., Park J., Tran H., Hu L.S., Hemmings B.A., Greenberg M.E.;
 RA "Protein kinase SGK mediates survival signals by phosphorylating the
 RA forkhead transcription factor FKHRL1 (FOXO3a).";
 RA Mol. Cell. Biol. 21:952-965(2001).
 RL [9]
 RL FUNCTION.
 RP MEDLINE-22284526; PubMed-12397388;
 RA Gamber N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
 RA Huber S.M., Kobayashi T., Cohen P.;
 RA "K(+)-channel activation by all three isoforms of serum- and
 RA glucocorticoid-dependent protein kinase SGK";
 RA Pfluegers Arch. 445:60-65(2002).
 CC -!- FUNCTION: protein kinase that plays an important role in cellular
 CC stress response. Activates certain potassium, sodium, and chloride
 CC channels, suggesting an involvement in the regulation of processes
 CC such as cell survival, neuronal excitability, and renal sodium
 CC excretion. Sustained high levels and activity may contribute to
 CC conditions such as hypertension and diabetic nephropathy. Mediates
 CC cell survival signals, phosphorylates and negatively regulates
 CC pro-apoptotic FOXO3a.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 CC phosphorylation.
 CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 CC in the pancreas, followed by placenta, kidney and lung.
 CC -!- INDUCTION: By serum and/or glucocorticoids. By excessive
 CC extracellular glucose and by TGF-beta, in cultured cells.
 CC -!- PTM: Regulated by phosphorylation. Phosphoinositide 3-kinase (PI3-
 CC kinase) pathway promotes phosphorylation at Ser-422 which in turn
 CC increases the phosphorylation of Thr-256 by PDK1.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL: Y10032; CRA71138.1;
 DR EMBL: AJ000512; CAA04146.1;
 DR EMBL: AF153609; AAD41091.1;
 DR EMBL: BC001263; AAH01263.1;
 DR HSP: P00317; 1YDR.
 DR Genew: HGNC:10810; SGK.
 DR MIM: 602938;
 DR GO: GO:0004674; P:protein serine/threonine kinase activity; TAS.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO: GO:0006950; P:response to stress; TAS.
 DR GO: GO:0006814; P:sodium ion transport; TAS.
 DR InterPro: IPR000961; Pkinase_C.
 DR InterPro: IPR000719; Prot_Kinase.
 DR InterPro: IPR0002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00433; pkinase_C; 1.
 DR ProDom: PD000001; Prot_Kinase; 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Apoptosis; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 98 355 PROTEIN KINASE.
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 236 236 PHOSPHORYLATION (BY PDK1).
 FT DOMAIN 131 141 LYS/GLU-RICH.
 FT MUTAGEN 256 256 T->A: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->D: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->E: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->A: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->D: 10-FOLD ACTIVATION.
 FT CONFLICT 381 381 E -> D (IN REF. 3 AND 4).
 SQ SEQUENCE 431 AA; 48956 MW; F3697DA57073399D CRC64;
 Query Match 100.0%; Score 2270; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 4.6e-157;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTVKTEAAGTUTYSGMKGWAILIAFMKQRRMGLNDFQKIANNSYACKHPEVQSILK1 60
 DB 1 MTVKTEAAGTUTYSGMKGWAILIAFMKQRRMGLNDFQKIANNSYACKHPEVQSILK1 60
 QY 61 SOPEPELMNANPSPPSPSQQINIGPSSNPHAKPSDFELKVIKGSFGKVLAKHKAEE 120
 DB 61 SOPEPELMNANPSPPSPSQQINIGPSSNPHAKPSDFELKVIKGSFGKVLAKHKAEE 120
 QY 121 EYFYAVKVLQKAILKAKKEEKHIMSEKRNVLNKKVHPFLVGLHFSFOTADKILYFVLIYIN 180
 DB 121 EYFYAVKVLQKAILKAKKEEKHIMSEKRNVLNKKVHPFLVGLHFSFOTADKILYFVLIYIN 180
 QY 181 GGELEVHLORECFLEPRARFAAEIASALGVLSNIVIRDLKPEINLLDSQGHIVTD 240
 DB 181 GGELEVHLORECFLEPRARFAAEIASALGVLSNIVIRDLKPEINLLDSQGHIVTD 240
 QY 241 FGICKENIEHNSTTSTFCGTPEYLAPEVLHKKOPYDRTVMWMLGAVIYEMLYGLPPFYSR 300
 DB 241 FGICKENIEHNSTTSTFCGTPEYLAPEVLHKKOPYDRTVMWMLGAVIYEMLYGLPPFYSR 300
 QY 301 NTAEMVDNLTNLKPLQKPNITNSARHLLGGLQKORTKLGAKDQDFMEIKSHVFFSLINW 360
 DB 301 NTAEMVDNLTNLKPLQKPNITNSARHLLGGLQKORTKLGAKDQDFMEIKSHVFFSLINW 360
 QY 361 DDLINKKITPPNPVNSGPNELRHDFDPEETPEVPNSIGKSPDSVLVTASVKEAAEFLG 420
 DB 361 DDLINKKITPPNPVNSGPNELRHDFDPEETPEVPNSIGKSPDSVLVTASVKEAAEFLG 420
 QY 421 FSYAPPTDSFL 431

```

Db          421 FSYAPPTDSFL 431
|||||
RESULT 2
SGK1_RABIT
ID          SGK1_RABIT      STANDARD;      PRT: 431 AA.
AC          Q9XT18;
DT          16-OCT-2001 (Rel. 40, Created)
DT          16-OCT-2001 (Rel. 40, Last sequence update)
DT          28-FEB-2003 (Rel. 41, Last annotation update)
DE          Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
DE          (Serum/glucocorticoid-regulated kinase 1)..
GN          SGK OR SGK1.
OS          Oryctolagus cuniculus (Rabbit).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX          NCBI_TaxID=9986;
RN          [1]
RS          SEQUENCE FROM N.A.
RC          STRAIN-New Zealand white;
RC          MEDLINE-99287894; PubMed-10358046;
RA          Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
RA          Fejes-Toth G.;
RT          "sgk is an aldosterone-induced kinase in the renal collecting duct.
RT          Effects on epithelial Na+ channels."
RL          J. Biol. Chem. 274:16973-16978(1999).
CC          -!- FUNCTION: Protein kinase that plays an important role in
CC          activating certain potassium, sodium, and chloride channels,
CC          suggesting an involvement in the regulation of processes such as
CC          cell survival, neuronal excitability, and renal sodium excretion.
CC          May be a key component of cellular stress response (By
CC          similarity).
CC          -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
CC          -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
CC          phosphorylation (By similarity).
CC          -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC          -----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
CC          -----
CC          EMBL; AF139639; AAD43303.1; -.
CC          RSP; P00517; IYDR.
CC          InterPro: IPR000961; Pkinase_C.
CC          InterPro: IPR000719; Prot_kinase.
CC          InterPro: IPR002290; Ser_thr_pkinase.
CC          Pfam: PF00069; pkinase; 1.
CC          Pfam: PF00433; pkinase_C; 1.
CC          ProDom: PD000001; Prot_kinase; 1.
CC          SMART: SM00133; S_TK_X; 1.
CC          SMART: SM00220; S_TK_C; 1.
CC          PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC          PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC          PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC          Apoptosis; Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC          phosphorylation.
CC          DOMAIN          98      355      PROTEIN KINASE.
CC          FT          NP_BIND          104      112      ATP (BY SIMILARITY).
CC          FT          BINDING          127      127      ATP (BY SIMILARITY).
CC          FT          ACT_SITE          222      222      BY SIMILARITY.
CC          FT          MOD_RES          256      256      PHOSPHORYLATION (BY PDPK1) (BY
CC          FT          SIMILARITY).
CC          SEQUENCE          431 AA; 48999 MW; 354898A77E8E38FD CRC64;
Query Match          97.6%; Score 2215; DB 1; Length 431;
Best Local Similarity 96.8%; Pred. No. 4.4e-153;
Matches 417; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

```

RESULT 3

SGK1_MOUSE

```

ID          SGK1_MOUSE      STANDARD;      PRT: 431 AA.
AC          Q9WVC6;
DT          16-OCT-2001 (Rel. 40, Created)
DT          16-OCT-2001 (Rel. 40, Last sequence update)
DT          15-SEP-2003 (Rel. 42, Last annotation update)
DE          Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
DE          (Serum/glucocorticoid-regulated kinase 1).
GN          SGK OR SGK1.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_TaxID=10090;
RN          [1]
RS          SEQUENCE FROM N.A.
RC          MEDLINE-99287894; PubMed-10358046;
RA          Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
RA          Fejes-Toth G.;
RT          "sgk is an aldosterone-induced kinase in the renal collecting duct.
RT          Effects on epithelial Na+ channels."
RL          J. Biol. Chem. 274:16973-16978(1999).
CC          [2]
CC          SEQUENCE FROM N.A.
CC          MEDLINE-20215285; PubMed-10751222;
CC          Shigaev A., Asher C., Latter H., Carly H., Reuveny E.;
CC          "Regulation of sgk by aldosterone and its effects on the epithelial
CC          Na(+) channel."
CC          Am. J. Physiol. 278:F613-F619(2000).
CC          [3]
CC          SEQUENCE FROM N.A.
CC          STRAIN-FVB/N;
CC          MEDLINE-2238257; PubMed-12477932;
CC          Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC          Klausner R.D., Collins F.S., Wagner J., Shenmen C.M., Schuler G.D.,
CC          Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC          Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC          Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC          Stapleton M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,
CC          Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,

```


RA Richards J.S., Fitzpatrick S.L., Clemens J.W., Morris J.K.,
 RA Alliston T., Sirotis J.;
 RA "Ovarian cell differentiation: a cascade of multiple hormones,
 RA cellular signals, and regulated genes";
 RA Recent Prog. Horm. Res. 50:223-254(1995).
 RN [4];
 RP INDUCTION BY P53.
 RC TISSUE-Mammary epithelium;
 RX MEDLINE-96218163; PubMed-8647846;
 RA Majvar A.C., Huang A.J., Phu P.T., Cha H.H., Firestone G.L.;
 RT "p53 stimulates promoter activity of the sgk.
 RT serum/glucocorticoid-inducible serine/threonine protein kinase gene
 RT in rodent mammary epithelial cells.";
 RN J. Biol. Chem. 271:12414-12422(1996).
 RN [5];
 RP PHOSPHORYLATION ON THR-256 BY PDPK1.
 RX MEDLINE-99286226; PubMed-10357815;
 RA Park J., Leong M.L., Buse P., Malyar A.C., Firestone G.L.,
 RA Hemmings B.A.;
 RT "Serum and glucocorticoid-inducible kinase (SGK) is a target of the p1
 RT 3-kinase-stimulated signaling pathway";
 RN EMBO J. 18:3024-3033(1999).
 CC -i- FUNCTION: Protein kinase that plays an important role in
 CC activating certain potassium, sodium, and chloride channels,
 CC suggesting an involvement in the regulation of processes such as
 CC cell survival, neuronal excitability, and renal sodium excretion.
 CC May be a key component of cellular stress response (By
 CC similarity). May also play an important role in the development of
 CC particular groups of neurons in the postnatal brain.
 CC -i- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -i- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 CC phosphorylation (By similarity).
 CC -i- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 CC in the ovary, thymus and lung.
 CC -i- INDUCTION: By dexamethasone and serum. By tumor suppressor p53 in
 CC mammary epithelial tumor cells. By FSH in granulosa cells. By
 CC injury to the central nervous system.
 CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L01624; AAA42137.1; -;
 CC HSSP: P00517; IYDR.
 CC InterPro: IPR000961; Pkinase_C.
 CC InterPro: IPR007019; Prot_kinase.
 CC InterPro: IPR002290; Ser_thr_pkinase.
 CC Pfam: PF00059; pkinase_1.
 CC Pfam: PF00433; pkinase_C_1.
 CC ProDom: PD000001; Prot_kinase_1.
 CC SMART: SM00133; S_TK_X: 1.
 CC SMART: SM00220; S_TK: 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST: 1.
 CC PROSITE: PS00011; PROTEIN_KINASE_DOM: 1.
 CC Apoptosis; Transferrase; Serine/threonine-protein kinase; ATP-binding;
 CC phosphorylation.
 CC DOMAIN 98 354 PROTEIN KINASE.
 CC FT BIND 104 112 ATP (BY SIMILARITY).
 CC FT BINDING 127 127 ATP (BY SIMILARITY).
 CC FT ACT_SITE 222 222 BY SIMILARITY.
 CC FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1).
 CC SQ SEQUENCE 430 AA; 48927 MW; 0D5845B04156F26D CRC64;
 CC -----
 CC Query Match 97.18; Score 2204.5; DB 1; Length 430;
 CC Best Local Similarity 96.8%; Pred. No. 2.5e-152;
 CC Matches 417; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 MTVKTEAAAGKTLTYSRMKGMVAIIIAFMKORRGLNDFIQKTANNKYAKKIPVQSLIKI 60
 DB 1 MTVKTEAAKSTLTYSRMKGMVAIIIAFMKORRGLNDFIQKLANNKYACKHPVQSLIKI 60
 QY 61 SQQPEPELMNANPSPSPSPSQINLGSPSSNPHAKPSDFHFLKVIKSGSPKVIILAKHKA 120
 DB 61 SQQPEPELMNANPSPSPSPSQINLGSPSSNPHAKPSDFHFLKVIKSGSPKVIILAKHKA 120
 QY 121 EYFYAVKVLQKKAIIKAKKEKHIMSEKRNVLKVKHPFLVGLHFSQTALUKLYEVLIDYIN 180
 DB 121 EAYAVKVLQKKAIIKAKKEKHIMSEKRNVLKVKHPFLVGLHFSQTADUKLYEVLIDYIN 180
 QY 181 GGEFLYHLORECFLEPRARFYAAEIASALGYLISLNIIVYRDLKPNILIDNSQHIVLID 240
 DB 181 GGEFLYHLORECFLEPRARFYAAEIASALGYLISLNIIVYRDLKPNILIDNSQHIVLID 240
 QY 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKLQDYDTVMWMLGAVLYEMLYGLPFPYSR 400
 DB 241 FGLCKENIEHNGTSTFCGTPEYLAPEVLHKLQDYDTVMWMLGAVLYEMLYGLPFPYSR 400
 QY 301 NTAEMVDNLIKPLQKPNITNSARHILLEGILQKDKTKRIGAKDPMELKSHVFESEINW 360
 DB 301 NTAEMVDNLIKPLQK - NITNSARHILLEGILQKDKTKRIGAKDPMELKSHIFFSILINW 359
 QY 361 DLLINKKITPPENPNVSGPNELRHDPETEEPPNNSIGKSPDSVLVTASVKEAAEAFIA 420
 DB 360 DLLINKKITPPENPNVSGPDLAHFOPEETEEPPVSSIGRSPDSILVTASVKEAAEAFIA 419
 QY 421 FSVAPPTDSFL 431
 DB 420 FSVAPPMDSFL 430
 RESULT 5
 SGK3_MOUSE
 ID SGK3_MOUSE STANDARD; PWT; 49% AA.
 AC Q9ERS3;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine/threonine-protein kinase Skk3 (EC 2.7.1.37)
 DE (Serum/glucocorticoid regulated kinase 3) (Serum/glucocorticoid
 DE regulated kinase-like) (Cytokine independent survival kinase).
 GN SKK3 OR SKK3 OR CISK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND MUTAGENESIS OF LYS-191.
 RX MEDLINE-20504817; PubMed-11050496;
 RA Liu D., Yang X., Songyang Z.;
 RT "Identification of CISK, a new member of the SGK kinase family that
 RT promotes IL-3-dependent survival";
 RL Curr. Biol. 10:1233-1236(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN-C57BL/6J; TISSUE-Porelimb, Ovary, and Uterus;
 RX MEDLINE-22354683; PubMed-12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaado I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Tagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kenzieski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Liersch B., Lyons P.A.,
 RA Maglott D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Kobayashi T., Deak M., Morrice N., Cohen P.;
 RT "Characterization of the structure and regulation of two novel
 RT isoforms of serum- and glucocorticoid-induced protein kinase.";
 RL Biochem. J. 344:189-197(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20054360; PubMed=10585774;
 RA Dai F., Yu L., He H., Zhao Y., Yang J., Zhang X., Zhao S.;
 RT "Cloning and mapping of a novel human Serum/Glucocorticoid regulated
 RT kinase-like gene, SGK1, to chromosome 8q12.3-q13.1.";
 RL Genomics 62:95-97(1999).
 RN [3]
 RP REVISIONS.
 RA Zhao Y.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Breast;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
 RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whitting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grummond J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=22284526; PubMed=12397388;
 RA Gamber N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
 RA Huber S.M., Kobayashi T., Cohen P., Lang F.;
 RT "K(+)-channel activation by all three isoforms of serum- and
 RT glucocorticoid-dependent protein kinase SGK";
 RL Pflügers Arch. 445:60-66(2002).
 CC -!- FUNCTION: Involved in the activation of potassium channels.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Localized in vesicle-like structures and in
 CC the early endosome (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 CC in pancreas, kidney liver, heart and brain and lower levels in
 CC lung, placenta and skeletal muscle.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.tsb-sib.ch/announce/>
 CC or send an email to license@tsb-sib.ch).
 CC -----
 CC EMBL: AF169035; AAF12758.1; ALT_INIT.
 CC EMBL: AF085233; AAF17031.2; -.
 CC EMBL: BC015326; AAF15326.1; -.
 CC GenBank: HGNC:10812; SGK1.
 CC MIM: 607591; -.
 CC InterPro: IPR000961; Pkinase.C.
 CC InterPro: IPR000719; Prot_Kinase.

DR InterPro: IPR001683; PX.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00433; pkinase_C; 1.
 DR Pfam: PF00787; PX; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00312; PX; 1.
 DR SMART: SM00133; S-TK_X; 1.
 DR SMART: SM00220; S-TK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00195; PX; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 12 124 PX.
 FT NP_BIND 162 419 PROTEIN KINASE.
 FT BINDING 168 176 ATP (BY SIMILARITY).
 FT ACT_SITE 191 191 ATP (BY SIMILARITY).
 FT MOD_RES 286 286 BY SIMILARITY.
 FT MUTAGEN 320 320 PHOSPHORYLATION (BY PDK1).
 FT CONFLICT 486 486 S->D: INCREASED ACTIVATION.
 FT CONFLICT 54 54 F -> V (IN REF. 2 AND 3).
 FT CONFLICT 127 127 D -> G (IN REF. 2 AND 3).
 FT CONFLICT 187 187 F -> V (IN REF. 2 AND 3).
 FT CONFLICT 294 294 L -> V (IN REF. 2 AND 3).
 FT CONFLICT 387 387 T -> R (IN REF. 2 AND 3).
 SQ SEQUENCE 496 AA: 57108 MW: 76A6CCEB69006CF1 CRC64;
 Query Match 64.8%; Score 1472; DB 1; Length 496;
 Best Local Similarity 67.2%; Pred. No. 3.le-99;
 Matches 279; Conservative 53; Mismatches 69; Indels 14; Gaps 3;
 QY 27 FKKQRMGLNDFIQKIANNKYACQHPQVQSIKISQOPELMNANSPSP----- 77
 DB 86 FIKQRRAGLNEFIQNLVRYPELYNHPDVRAPLQMDSPKH----QSDPEDEDEKSSAKLH 141
 QY 78 SPSSQINLGPSNPHAKYSDUHFELKVIQKSGFGKVLARHAKAEVYFAVKVQKAILKK 137
 DB 142 STSQINLGPSNPHAKPTDFELKVIQKSGFGKVLAKRKLQDGFYAVKVLQKKVLNR 201
 QY 138 KEEKIMSPERNVLLKNRPFVINGLHFSQFADKILYFVLDYINGSELFYHQRKCFLEP 197
 DB 202 KQKQHMAERNVLLKNRPFVINGLHFSQFADKILYFVLDYINGSELFYHQRKCFLEP 261
 QY 198 RARFYAAETASALGYLHSLNIVYRDLKPENILLDSQGHVLTDFGLCKENETSTFT 257
 DB 262 RARFYAAETASALGYLHSLIKIYRDLKPENILLDSQGHVLTDFGLCKEGIALSDITTF 321
 QY 258 CGTPEYLAPEVLHKKOPYDRTVDMWCLGAVLYEMLYGLPPYSRNTAEYDNLNKKPLQK 317
 DB 322 CGTPEYLAPEVIRKOPYDNTVDMWCLGAVLYEMLYGLPPYCRDVAEYDNLNKKPLSLR 381
 QY 318 PNITSARHLEGLQKDKTRKLGAKDDFMELKSHVFFSLINWDDLNKKITPPFPNPVS 377
 DB 382 PGVSLTAWLSILEELKDRQNLGAKEDFLEIQNHPPFESLSWADLVQKKIPPPFPNVA 441
 QY 378 GNEILRHDPETEPVPNSITGKSPDSVLYTASYKEAAEAFGLGFSYAPPT-DSFL 431
 DB 442 GPDIDRNFTAFTEETVPYSVCVSDYSIVNASVLEADDAFVGFSYAPPSEHLFV 496
 RESULT 7
 SGK2_HUMAN
 ID SGK2_HUMAN STANDARD; PRT; 427 AA.
 AC Q9HBY8; Q9HKG6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.37)
 DE (Serum/glucocorticoid regulated kinase 2).
 GN SGK2.
 OS Homo sapiens (Human).

QY 132 KALKKKKEKHIMSERVLLKNVHPFLVGLHFSFOTADKLYFVLDYINGGELFYHLQRE 191
 DB 69 KSLKKNQNHIMAEIRVLLKNVHPFLVGLHFSFOTADKLYFVLDYINGGELFYHLQRE 128
 QY 192 RCFLEPRARFYAAEIASALGYLHSLNIVYRDLKPNILLOSOGHVLVDFGLCKENIEHN 251
 DB 129 RFELEPRARFYAAEIASALGYLHSLNIVYRDLKPNILLOSOGHVLVDFGLCKENIEHN 188
 QY 252 STTSTFCGPEYLAPVLAPELVHKOYDRTVDMWCLGAVLYEMLYGLPPFYGRNTAEMYNILN 311
 DB 189 ETTSTFCGPEYLAPVLAPELVHKOYDRTVDMWCLGAVLYEMLYGLPPFYGRNTAEMYNILN 248
 QY 312 KPLQLEPNTNSARHLLGLLOKORTKRLGAKDDMEIKSHVFFSLINWDDLKINKITPP 371
 DB 249 OPLQIFGGRFYAACDQLGLLQKQORGLSKSDEFLDKNHFSPINWDDLYHRLRTPP 308
 QY 372 FPNVSGPNELRHFDEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGFSYAPPTDSFL 431
 DB 309 FPNVSGPADLKHEDPEFTQEAHSVKSIGCTPTDV---ASSSGASSAFILGFSYAQDDDDIL 365

RESULT 9

SGK2_RAT STANDARD; PRT; 302 AA.
 AC Q8R4U9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serine/threonine-protein kinase SGK2 (EC 2.7.1.37)
 DE (Serum/glucocorticoid regulated kinase 2) (Fragment).
 GN SGK2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID-10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Kidney;
 RA Feng Y.X., Huber S.M., Waernstges S., Lang F.;
 RA "SGK2 and SGK3 mRNA expression in rat kidney";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Involved in the activation of potassium channels (by similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF361756; AAL91351.1;
 DR InterPro; IPR000719; Prot.kinase.
 DR InterPro; IPR002290; Ser.thr.pkinase.
 DR Pfam; PF00069; pkinase; 1
 DR Pfam; PF00433; pkinase_C; 1.
 DR PRODOM; PD000001; Prot.Kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_NTP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.

FT NON_TER 1 1
 FT DOMAIN 18 275
 FT BIND 24 32
 FT BINDING 47 47
 FT ACT_SITE 142 142
 FT MOD_RES 176 176

PROTEIN KINASE.
 ATP (BY SIMILARITY).
 ATP (BY SIMILARITY).
 BY SIMILARITY.
 PHOSPHORYLATION (BY PDK1) (BY

FT NON_TER 302 302
 SQ SEQUENCE 302 AA; 34622 MW; A74EE3F424383D66 CRC64;
 SIMILARITY).
 Query Match 54.3%; Score 1232; DB 1; Length 302;
 Best Local Similarity 73.7%; Pred. No. 3-8e-82;
 Matches 221; Conservative 46; Mismatches 33; Indels 0; Gaps 0;
 QY 83 INLPSSNPHAKYSDFFHFKVIGKSGFKVLLARHAAEEFYAVKVLQKAIILKKKEPKH 142
 DB 3 INLPSSNPHAKYSDFFHFKVIGKSGFKVLLARHAAEEFYAVKVLQKAIILKKKEPKH 62
 QY 143 IMSENRVLLKNVHPFLVGLHFSFOTADKLYFVLDYINGGELFYHLQREKCFLEPRARFY 202
 DB 63 IMSENRVLLKNVHPFLVGLHFSFOTADKLYFVLDYINGGELFYHLQREKCFLEPRARFY 122
 QY 203 AAEIASALGYLHSLNIVYRDLKPNILLOSOGHVLVDFGLCKENIEHNSTTSTFCGPE 262
 DB 123 TAESAASALGYLHSLNIVYRDLKPNILLOSOGHVLVDFGLCKENIEHNSTTSTFCGPE 182
 QY 263 YLAPEVLHKOYDRTVDMWCLGAVLYEMLYGLPPFYGRNTAEMYNILNKKPKIKNITN 322
 DB 183 YLAPEVLHKOYDRTVDMWCLGAVLYEMLYGLPPFYGRNTAEMYNILNKKPKIKNITN 242
 QY 323 SARHLLGLLOKORTKRLGAKDDMEIKSHVFFSLINWDDLKINKITPPNPNVSGPNEL 482
 DB 243 AACDLQGLLHKKQORGLSKSDEFLDKNHFSPINWDDLYHRLRTPPNNVSGPADL 302

RESULT 10

AKT3_HUMAN STANDARD; PRT; 479 AA.
 AC Q9V243; Q96QV3; Q9UPP5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE RAC-gamma serine/threonine protein kinase (EC 2.7.1.-) (RAC-PK-gamma)
 DE (Protein Kinase Akt-3) (Protein Kinase B, gamma) (PKB gamma) (STK-2).
 GN AKT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID-9606;
 RN [1]
 RC SEQUENCE FROM N.A., AND MUTAGENESIS.
 RC MEDLINE-99194749; PubMed-10092583;
 RA Brodebeck D., Cron P., Hemmings B.A.;
 RT "A human protein kinase B gamma with regulatory phosphorylation sites in the activation loop and in the C-terminal hydrophobic domain";
 RL J. Biol. Chem. 274:9133-9136(1999).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC MEDLINE-99225329; PubMed-10208883;
 RA Nakatani K., Sakae H., Thompson D.A., Weigel R.J., Roth R.A.;
 RT "Identification of a human Akt3 (protein kinase B gamma) which contains the regulatory serine phosphorylation site";
 RL Biochem. Biophys. Res. Commun. 257:906-910(1999).
 RN [3]
 RC SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RC MEDLINE-99421751; PubMed-10491192;
 RA Masure S., Haefner B., Wesseling J.-J., Hoeftnagel E., Mortier E., Verhaesselt P., Tuytelaars A., Gordon R., Richardson A.;
 RT "Molecular cloning, expression and characterization of the human serine/threonine kinase Akt-3";
 RL Eur. J. Biochem. 265:353-360(1999).
 RN [4]
 RC SEQUENCE FROM N.A. (ISOFORM 1).
 RA Li X., Yu L., Huang H., Zhang M., Zhao Y., Zhao S.;
 RT "Cloning of a novel human cDNA, STK-2, which encodes a rat serine-threonine protein kinase (STK) homolog";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]

RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE-testis;
RX MEDLINE-21154917; PubMed-11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Mambuti R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs";
RL Genome Res. 11:422-435(2001).
[6]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF THR-305 AND
RP THR-447.
RX PubMed-11387345;
RA Brodbeck D., Hill M.M., Hemmings B.A.;
RT "two splice variants of PKB gamma have different regulatory capacity
RT depending on the presence or absence of the regulatory phosphorylation
RT site Ser-472 in the C-terminal hydrophobic domain";
RL J. Biol. Chem. 276:29550-29558(2001).
CC -!- FUNCTION: IGF-1 LEADS TO THE ACTIVATION OF AKT3, WHICH MAY PLAY A
CC ROLE IN REGULATING CELL SURVIVAL. CAPABLE OF PHOSPHORYLATING
CC SEVERAL KNOWN PROTEINS. TRUNCATED ISOFORM 2/PKB GAMMA 1 WITHOUT
CC THE SECOND SERINE PHOSPHORYLATION SITE COULD STILL BE STIMULATED
CC BUT IN A LESSER EXTENT.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED AFTER
CC CELL STIMULATION LEADING TO ITS TRANSLOCATION.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=PKB gamma;
CC IsoId=Q9Y243-1; Sequence=Displayed;
CC Name=2; Synonyms=PKB gamma.1;
CC IsoId=Q9Y243-2; Sequence=VSP_004947;
CC -!- TISSUE SPECIFICITY: IN ADULT TISSUES, IT IS HIGHLY EXPRESSED IN
CC BRAIN, LUNG AND KIDNEY, BUT WEAKLY IN HEART, TESTIS AND LIVER. IN
CC FETAL TISSUES, IT IS HIGHLY EXPRESSED IN HEART, LIVER AND BRAIN
CC AND NOT AT ALL IN KIDNEY.
CC -!- DOMAIN: BINDING OF THE PH DOMAIN TO THE PHOSPHATIDYLINOSITOL 3-
CC KINASE ALPHA (PI(3)K) RESULTS IN ITS TARGETING TO THE PLASMA
CC MEMBRANE.
CC -!- PTM: PHOSPHORYLATED ON THREONINE AND SERINE RESIDUES.
CC PHOSPHORYLATION ON BOTH SITES IS REQUIRED FOR FULL ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. RAC
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 PH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF1244141; A029089.1; .
CC EMBL: AF135794; A024196.1; .
CC EMBL: AF085234; A040392.1; .
CC EMBL: AJ245709; CAB53537.1; .
CC EMBL: AL117525; CAB55977.1; ALT_TERM.
CC EMBL: AY005799; AAF91073.1; .
CC EMBL: AY5380; A59380.
CC HSP: P05132; IFMO.
CC Genew: HGNC:393; AKT3.
CC GO: GO:0006468; P:protein amino acid phosphorylation; TAS.
CC GO: GO:0007165; P:signal transduction; TAS.
CC InterPro: IPR001849; PH.
CC InterPro: IPR000961; Kinase_C.
CC InterPro: IPR000719; Ser_kinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC Pfam: PF00169; PH; 1.
CC Pfam: PF00069; pkinase; 1.
CC Pfam: PF00433; pkinase_C; 1.

DR PRINTS: PR00109; TYRKINASE.
DR Prodom: PD000001; Prot_kinase; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00133; S_TK_X; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00003; PH_DOMAIN; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 5 107
FT NP_BIND 148 405
FT BINDING 154 162
FT BINDING 177 177
FT ACT_SITE 271 271
FT MOD_RES 305 305
FT MOD_RES 472 472
FT VARSPLIC 452 479
FT MUTAGEN 305 305
FT MUTAGEN 305 305
FT MUTAGEN 447 447
FT MUTAGEN 447 447
FT MUTAGEN 472 472
FT MUTAGEN 472 472
FT SEQUENCE 479 AA; 55774 MW; F08BDD6502E78FB CRC64;
Query Match 40.0%; Score 909; DB 1; Length 479;
Best Local Similarity 45.6%; Pred. No. 1.6e-58;
Matches 191; Conservative 65; Mismatches 135; Indels 28; Gaps 7;
QY 28 MKQRRLNDFIQK-----IANNYSACKHPE-----VQSILKISQPOEPEIMANP 73
DB 62 MKTERPKNTRIRCLQNTTIVERTFVDTPEREEMTEAIOAVADRLOROEENMNCSP 121
QY 74 SPSPSPSQINLG---PSSNPHAK---PSDFHLKVLKCGSPGKVLARKHAEVYAV 126
DB 122 T-----SQIDNIGEEEMDASTTHHKRTKMTNDFUYLLKLGKTFGKVLVLRKAS:KYIAM 176
QY 127 KVLQKAILKKKEEKHKIMSENVLLKNVYKHPFLVGLHPSFQTADKIYFVLDTNGGEIFY 186
DB 177 KILKEVILIAKPEVAHTLTESRV-LKNTRHPELTSLKYSFQTKDRICFVMEVNGELFF 235
QY 187 HQRECFLEPRARYAAEIASALGYLHSLNIVYRDLPENILDSQGHIVLTVFGLCKE 246
DB 236 HLSRERVFSEDTREYGAETVSDYLSGKIVSRDLKLENLMDKDGHIKITDFGLCKE 295
QY 247 NIEHNSSTTFCGTEPYLAPEVLHQPDRVDMWCLCAVLYEMLYGLPPYFNTADMY 306
DB 296 GITDAATMKTTCGTEPYLAPEVLENDNDYGRAVDWNGLVGVVYEMMCGELPYNQDHEKLP 355
QY 307 DNILNKPLOLKNITNSARHLLEGLLQDKDRKRL-GAKDDFMEIKSHVFFSLINWDLIN 365
DB 356 ELIMEDIKFRPTLSSDAKSLISGLLIKDPNKRLOGGPDDAKEINRHSFFSGVANNQVDYD 415
QY 366 KKITPPFNPNVSGPNELRHDFEPTTEEPVNPNSJCKSPDVLVTASVKEAAEAFIGSYA 424
DB 416 KALVPPFKQVTSYTDTRFDSEFTRQITITTPPEKYDEGDMCDMNERPHFPQFSYS 474
RESULT 11
ID AKT3_MOUSE
AC Q9W0A6; STANDARD; PRT: 479 AA.

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE RAC-gamma serine/threonine protein kinase (EC 2.7.1.1-) (RAC-PK-gamma)
 DE (Protein kinase Akt-3) (Protein kinase B, gamma) (PKB gamma).
 GN AKT3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=99194749; PubMed=10092583;
 RA Brodebeck D., Cron P., Hemmings B.A.;
 RT "A human protein kinase B gamma with regulatory phosphorylation sites
 RT in the activation loop and in the C-terminal hydrophobic domain.";
 RL J. Biol. Chem. 274:9133-9136(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brain;
 RX PubMed=11387345;
 RA Brodebeck D., Hill M.M., Hemmings B.A.;
 RT "Two splice variants of PKB gamma have different regulatory capacity
 RT depending on the presence or absence of the regulatory phosphorylation
 RT site Ser-72 in the C-terminal hydrophobic domain.";
 RL J. Biol. Chem. 276:29550-29558(2001).
 CC -!- FUNCTION: IGF-1 LEAD TO THE ACTIVATION OF AKT3, WHICH MAY PLAY A
 CC ROLE IN REGULATING CELL SURVIVAL. CAPABLE OF PHOSPHORYLATING
 CC SEVERAL KNOWN PROTEINS. TRUNCATED ISOFORM 2/PKB GAMMA 1 WITHOUT
 CC THE SECOND SERINE PHOSPHORYLATION SITE COULD STILL BE STIMULATED
 CC BUT IN A LESSER EXTENT (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOSOLIC AND MEMBRANE-ASSOCIATED AFTER
 CC CELL STIMULATION LEADING TO ITS TRANSLOCATION.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms-2;
 CC Name-1: Synonyms=PKB gamma;
 CC IsoId-Q9MUA6-1; Sequence=Displayed;
 CC Name-2: Synonyms=PKB gamma 1;
 CC IsoId-Q9MUA6-2; Sequence=VSP_004948;
 CC -!- TISSUE SPECIFICITY: ISOFORM 1 IS EXPRESSED IN PROSTATE, TESTIS,
 CC UTERUS AND MAMMARY GLAND AND ISOFORM 2 IS EXPRESSED IN PROSTATE,
 CC TESTIS AND MAMMARY GLAND.
 CC -!- DOMAIN: BINDING OF THE PH DOMAIN TO THE PHOSPHATIDYLINOSITOL 3-
 CC KINASE ALPHA (PI(3)K) RESULTS IN ITS TARGETING TO THE PLASMA
 CC MEMBRANE.
 CC -!- PTM: PHOSPHORYLATED ON THREONINE AND SERINE RESIDUES.
 CC PHOSPHORYLATION ON BOTH SITES IS REQUIRED FOR FULL ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. RAC
 CC SUBFAMILY.
 CC -!- SIMILARITY: Contains 1 PH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@sib-sib.ch.
 CC -----
 DR EMBL; AF124142; AAD29090.1;
 DR HSSP; P05132; ICTP.
 DR MGD; MG1:1345147; Akt3.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase_C; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000001; Prot_kinase; 1.
 DR SMART; SM00233; PH; 1.

DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00003; PH_DOMAIN; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation; Alternative splicing.
 FT DOMAIN 5 107
 FT NP_BIND 148 405
 FT NP_BIND 154 162
 FT BINDING 177 177
 FT ACT_SITE 271 271
 FT MOD_RES 305 305
 FT MOD_RES 472 472
 FT VARSPPLIC 452 479
 FT FT
 FT FT
 SQ SEQUENCE 479 AA: 55714 MW: 508ACDF574388PB CRC64;
 Query Match 40.0%; Score 909; DB 1; Length 479;
 Best local Similarity 45.6%; Pred. No. 1.6e-58;
 Matches 191; Conservative 65; Mismatches 135; Indels 28; Gaps 7;
 QY 28 MKORRMGLNDFIOK-----IANSYACKHP-----VOSILKISQOPELMANP 73
 DB 62 MKTERPKNTEIIRCLQMTTIVTERTEHVDTPEEREETEAQAVADRLOREERMNCSP 121
 QY 74 SPSPSSQOINLG---PSSNPFAK---PSDFHLKVIQKSGFKVLLARHKAEEVFYAV 126
 DB 122 T-----SQIDNIGEEEMDASTTHHKRTMNDFDYLLKLGKGTFCVKLVREKASKYYAM 176
 QY 127 KVLQKAKILKKKEKHINSERNVLLKNYKIPFLVGLHFSQTDKLYLVLYINQGLIFY 186
 DB 177 KILKKEVIIAKDEVAHTLTSERV-LKNTRHPPELTSKYSFQTKDCLCEVNEVNGELFF 235
 QY 187 HLQRECFLEPRARFYAABIASALGYLSNLIVYRDLKPKENILLDSQGHIVLTDFGLCKE 246
 DB 236 HLSRERFVEDRTFYCAEIVSALDYLSHSGKIVYRDLKLENLMDKDGKIKITDFGLCKE 295
 QY 247 NIEHNTSTFEGCTPEYLAPEVLHKKQPYDRTVQWNCGLGAVLYENLYGLPPYKNTAEM 306
 DB 296 GITDAATMTFCCTPEYLAPEVLENDYGRVDMGLGVVYVMMCKRLPYNQJHKLK 355
 QY 307 DNILNPLQKPNITNSARHLLLEGLLQKQRTKRL-GAKDDFMEIKSHVFFSLINMDJIN 365
 DB 356 ELILMEDIKFPRTLSSDAKSLGLLIKDPNKRLLGGPDDAKEIMRHSEFFSGVNWQDVYD 415
 QY 366 KKITPPNPVNSGPNELRHFDEPPEPNSIGKSPSVLYTASYKAAAFALCPSYA 424
 DB 416 KKLVPFPKPOVTSETOTRYDFDEFTAOITITTPPEKYDDGDMGDMNRRPHFPQFSYS 474

RESULT 12

ID AKT3_RAT
 ID AKT3_RAT
 AC Q63484; STANDARD; PRT; 454 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE RAC-gamma serine/threonine protein kinase (EC 2.7.1.1-) (RAC-PK-gamma)
 DE (Protein kinase Akt-3) (Protein kinase B, gamma) (PKB gamma).
 GN AKT3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 EX MEDLINE=9603640; PubMed=7498143;
 RA Koniishi H., Kuroda S., Tanaka M., Matsuzaki H., Ono Y., Kameyama K.,
 RA Haga T., Kikkawa U.

236 HLSRERVSDRTFRFGAELVGSALDYLSHGKIVYRDLENLMLDKOCHIKITUGLAK 295

247 NIEHNSTSTFCQPEYLAPVHLKQPYDRTVDMWICIGAVLYEMLYGLPYPSKNTAEKY 306

296 GITDAAMTKTFCQPEYLAPVHLKQPYDRTVDMWICIGAVLYEMLYGLPYPSKNTAEKY 355

307 DNIILNKPQLQKPNITNSARHLEGLLOKDRKRL-GAKDDFMEIKSHVFSFLINWDDLIN 365

356 ELILMEDIKFPRLLSDAKSLGSLGLIKDPNKRKGQGGDDPKENRHSFFSCVNWQVYD 415

366 KKITPPNPVNSGPNELKHFDEFTPEP 394

416 KKLVPPEKQVTSSETDIRYFDEFTAQTI 444

RESULT 13

ID	KAKT_MLVAT	STANDARD:	PRT:	501 AA.
AC	P31748;			
DT	Q1-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	AKT kinase transforming protein (EC 2.7.1.-).			
GN	V-AKT.			
OS	AKT8 murine leukemia virus.			
OC	Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.			
OX	NCBI_TaxID=11790;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	BELLING-92022574; PubMed=1833819;			
RX	Bellacosa A., Testa J.R., Staal S.P., Tsichlis P.N.;			
RT	"A retroviral oncogene, akt, encoding a serine-threonine kinase			
RT	containing an SH2-like region.";			
RL	Science 254:274-277(1991).			
CC	-1- PTH. AUTOPHOSPHORYLATED ON THR AND SER RESIDUES.			
CC	-1- MISCELLANEOUS: THIS PROTEIN IS EXPRESSED AS A FUSED GAG-AKT			
CC	POLYPROTEIN.			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	RAC SUBFAMILY.			
CC	-1- SIMILARITY: Contains 1 PH domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/annouce/)			
CC	or send an email to license@isb-sib.ch .			
CC	-----			
CC	EMBL; M80675; AAA2545.1; -.			
DR	HSP; P05132; ICTP.			
DR	InterPro; IPR001849; PH.			
DR	InterPro; IPR000961; Pkinase_C.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002290; Ser_thr_pkinase.			
DR	Pfam; PF00169; PH; 1.			
DR	Pfam; PF00069; Pkinase; 1.			
DR	Pfam; PF00433; Pkinase_C; 1.			
DR	ProDom; P0000001; Prot_kinase; 1.			
DR	SMART; SM00233; PH; 1.			
DR	SMART; SM00133; S_TK_X; 1.			
DR	SMART; SM00220; S_TKC; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.			
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE; PS50003; PH_DOMAIN; 1.			
KW	Oncogene; Transferase; Serine/threonine-protein kinase;			
KW	Phosphorylation; ATP-binding.			
FT	DOMAIN 26 129 PH			
FT	DOMAIN 171 429 PROTEIN KINASE.			
FT	NP_BIND 177 185 ATP (BY SIMILARITY).			
FT	BINDING 200 200 ATP (BY SIMILARITY).			
FT	ACT_SITE 295 295 BY SIMILARITY.			

```

FT  MOD_RES 347 347 PHOSPHORYLATION (BY SIMILARITY).
SQ  SEQUENCE 501 AA; 57870 MW; 5AEFDE58CD42F773 CRC64;

Query Match 38.1%; Score 865; DB 1; Length 501;
Best Local Similarity 43.7%; Pred. No. 2.6e-55;
Matches 185; Conservative 72; Mismatches 130; Indels 36; Gaps 9;

QY 28 MKQRMGLNDFIQK-----IANNYSACKHPE-----VQSILKISQPEPELMNANP 73
DB 84 MKTERPRNTFIIRCLQWTTVIERTPHVETPEEREENATAITQTVADGLKQREETMDFRS 143
QY 74 SPPSPS---QQINLGPSSNPHAKP-----SDRFLKVIKGSFGKVLARHKAEEVF 123
DB 144 GSPSDNSGAEMEVSLSL-----AKPKHRTVMNEFFYLKLGKGFVKVILYKATGRY 196
QY 124 YAVVLQKAILKKKEEKHIMSERNVLLKNVHPFLVNGURFSQFADKLYFVLVDYINGGE 183
DB 197 YAMKILKEVIVAKDEVAHTLSE- NRVLQNSRHFFLTALKYSFQTHDRLCFVMEYANGE 255
QY 184 LFYHLQRECFLEPRARFYAAETASALGYLHS- LNTVYRLDLPENILLDSQGHVLTDFG 242
DB 256 LFYHLGRVFSDDRARFYGAETVSDYLHSEKNVYVYRDLEMLKLEMLKQGHVLTDFG 315
QY 243 LCKENIEHNTSTTFCGTEPEYLAPEVHLKQPDYDRTVDWCLGAVLYEMLYGLPFPYSRNT 302
DB 316 LCKEGIKDGTATKTCGTEPEYLAPEVLENDYDGRAVDWMLGVLGVVYVYEMMCGRLPEYNODH 375
QY 303 AEMVYLNKPLQKPNITNSARHLLGLELQKDRKRL- GAKDPMKEIKSHVFFSLINWD 361
DB 376 EKIFELLMEIEIRFPTLGEPAKSLGGLLKQDPTORLGGSEDAAKEIMQHREFANIVWQ 435
QY 362 DLINKKTTFPNVPNSGPNELRHPDTEEPVNSIGKSPDSVLVTASVKAEEAFILGF 421
DB 436 DVEYKLLSPKQVTSOTRYTDEFTAQMITITPPQDDSMCECVDS-ERRPHFPQF 493
QY 422 SYA 424
DB 494 SYS 496

RESULT 14
ID KRAC_MOUSE STANDARD; PRT; 480 AA.
AC P31750; 062274;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE RAC-alpha serine/threonine kinase (EC 2.7.1.-) (RAC-PK-alpha) (AKT1
DE kinase) (protein kinase B) (PKB) (C-AKT) (Thymoma viral proto-
DE oncogene).
GN AKT1 OR AKT OR RAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bousquets X., Powell C.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=AKR/J; TISSUE=Thymus;
RX MEDLINE=93173519; PubMed=8437858;
RA Bellacosa A., Franke T.F., Gonzalez-Portal M.B., Datta K., Taguchi T.,
RA Gardner J., Cheng J.Q., Testa J.R., Tschlis P.N.;
RT "Structure, expression and chromosomal mapping of c-akt: relationship
RT to v-akt and its implications.";
RL Oncogene 8:745-754(1993).
CC -!- FUNCTION: General protein kinase capable of phosphorylating
CC several known proteins.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear after activation by
CC integrin-linked protein kinase 1 (ILK1) (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed. Low levels found in liver

```

```

CC with slightly higher levels present in thymus and testis.
CC -!- DOMAIN: Binding of the ph domain to the phosphatidylinositol 3-
CC kinase alpha (PI3K) results in its targeting to the plasma
CC membrane.
CC -!- PTM: Phosphorylation on Thr-308, Ser-473 and Tyr-474 is required
CC for full activity.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: Contains 1 PH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M94335; AAA18254.1; -.
CC EMBL: X65687; CAA46620.1; -.
CC PIR: S33364; S33364.
CC HSP: P05132; ICTP.
CC MGD: MGI:87986; Akt1.
CC GO: GO:0005737; Cytoplasm; IDA.
CC GO: GO:0006915; P:apoptosis; IDA.
CC GO: GO:0007281; P:germ-cell development; IDA.
CC InterPro: IPR001849; PH.
CC InterPro: IPR000961; PKinase.C.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR002290; Ser_thr_kinase.
CC Pfam: PF00169; PH; 1.
CC Pfam: PF00069; pkinase; 1.
CC Pfam: PF00433; pkinase.C; 1.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00233; PH; 1.
CC SMART: SM00133; S_TK_X; 1.
CC SMART: SM00220; S_TK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00003; PH_DOMAIN; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation.
FT DOMAIN 5 108 PH.
FT DOMAIN 150 408 PROTEIN KINASE.
FT NP_BIND 156 164 ATP (BY SIMILARITY).
FT BINDING 179 179 ATP (BY SIMILARITY).
FT ACT_SITE 274 274 BY SIMILARITY.
FT MOD_RES 308 308 PHOSPHORYLATION (BY PDPK1) (BY
FT SIMILARITY).
FT MOD_RES 473 473 PHOSPHORYLATION (BY ILK1) (BY
FT SIMILARITY).
FT MOD_RES 474 474 PHOSPHORYLATION (BY SIMILARITY).
FT CONFLICT 367 367 A -> R (IN REF. 2).
SQ SEQUENCE 480 AA; 55622 MW; 18D21018593B5A98 CRC64;

Query Match 38.0%; Score 863; DB 1; Length 480;
Best Local Similarity 43.7%; Pred. No. 3.5e-55;
Matches 185; Conservative 71; Mismatches 131; Indels 36; Gaps 9;

QY 28 MKQRMGLNDFIQK-----IANNYSACKHPE-----VQSILKISQPEPELMNANP 73
DB 63 MKTERPRNTFIIRCLQWTTVIERTPHVETPEEREENATAITQTVADGLKQREETMDFRS 122
QY 74 SPPSPS---QQINLGPSSNPHAKP-----SDRFLKVIKGSFGKVLARHKAEEVF 123
DB 123 GSPSDNSGAEMEVSLSL-----AKPKHRTVMNEFFYLKLGKGFVKVILYKATGRY 175
QY 124 YAVVLQKAILKKKEEKHIMSERNVLLKNVHPFLVGLHFSFQADKLYFVLVDYINGGE 183
DB 176 YAMKILKEVIVAKDEVAHTLSE- NRVLQNSRHFFLTALKYSFQTHDRLCFVMEYANGE 234
QY 184 LFYHLQRECFLEPRARFYAAETASALGYLHS- LNTVYRLDLPENILLDSQGHVLTDFG 242

```

```

Db 235 LFFHLSRKFSEDRARYGAIVSALDYHSEKNVYRDVWKLKLEMLDKDGHKIKITDFG 294
Qy 243 LCKENIEHNSITSTFCGTPPEYLAPVHLKQPDYDRTVDWMLCGLAVLYEMLYGLPPPYSRNT 302
Db 295 LCKEGIKDGTWKTFCGTPPEYLAPVLEDNDYGRAVDWMLGVLVYVWYEMMCGRLPYNDQH 354
Qy 303 ARMYONILNKLPLKPNITNSARHLEGLLQKDKTKRL-GAKDDPMEIKSHVFFSLINWD 361
Db 355 ELKFLLEIMEEIAFRTLLGPEAKSLSLGLLKKDPTQRLGGGSEDAKEIMQHRFFANIVMQ 414
Qy 362 DLINKKITPPPNVSGPNELRHFPETEEVPNSICKSPDSVLVTASVKEAARAFGLGF 421
Db 415 DVEKSLSPFPKQVTSYDTRFYDEEFTAQMITITPPDQDSMECVDS--ERRPHFPQF 472
Qy 422 SYA 424
Db 473 SYS 475

RESULT 15
KRAC_RAT
ID KRAC_RAT STANDARD; PRT; 480 AA.
AC P47196;
Dt 01-FEB-1996 (Rel. 33, Created)
Dt 01-FEB-1996 (Rel. 33, Last sequence update)
Dt 28-FEB-2003 (Rel. 41, Last annotation update)
DE RAC-alpha serine/threonine kinase (EC 2.7.1.-) (RAC-PK-alpha)
DE (Protein kinase B) (PKB).
OS AKT1.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=95091823; PubMed=7999118;
RA Konishi H., Shinomura T., Kuroda S.I., Ono Y., Kikkawa U.;
RT "Molecular cloning of rat RAC protein kinase alpha and beta and their
RT association with protein kinase C zeta.";
RL Biochem. Biophys. Res. Commun. 205:817-825(1994).
CC -!- FUNCTION: General protein kinase capable of phosphorylating
CC several known proteins.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear after activation by
CC integrin-linked protein kinase 1 (ILK1) (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed. Low levels found in liver
CC with slightly higher levels present in thymus and testis.
CC -!- DOMAIN: Binding of the ph domain to the phosphatidylinositol 3-
CC kinase alpha (PI(3)K) results in its targeting to the plasma
CC membrane.
CC -!- PTM: Phosphorylation on Thr-308, Ser-473 and Tyr-474 is required
CC for full activity.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC RAC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 PH domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; D30040; BAA06279.1; -.
DR JC2437; JC2437.
DR HSSP; P05132; 1CTP.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; pkinase.C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00169; PH; 1.

```

```

DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase.C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00133; S-TK_X; 1.
DR SMART; SM00220; S-TRC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 5 108 PH.
FT DOMAIN 150 408 PROTEIN KINASE.
FT NP_BIND 156 164 ATP (BY SIMILARITY).
FT BINDING 179 179 ATP (BY SIMILARITY).
FT ACT_SITE 274 274 BY SIMILARITY.
FT MOD_RES 308 308 PHOSPHORYLATION (BY PDPK1) (BY
FT SIMILARITY).
FT MOD_RES 473 473 PHOSPHORYLATION (BY ILK1) (BY
FT SIMILARITY).
FT MOD_RES 474 474 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 480 AA; 55735 MW; 5DCAAE7134366D04 CRC64;

Query Match 38.0%; Score 863; DB 1; Length 480;
Best Local Similarity 43.8%; Pred. No. 3.5e-55;
Matches 182; Conservative 75; Mismatches 137; Indels 22; Gaps 8;

Qy 28 MKQRRMGLNDFIQK-----IANNYSACKHEP-----VQSILKISQPEPELMNANP 73
Db 63 MKTERPRNTFIIRCLQWITVIERTFHVETPEEREEMTTAQTADVADGLKROEETMDPFRS 122
Qy 74 SPPP--SPSQOINILGSSNPH-AKPSDFHLKVLGKSGFVKVLLARHKAEVEFYAVKVLQ 130
Db 123 GSPSDNSGAEKEVALAKPKRVMTMNEFEYKLLGKGTGKVLVKEKATRYTAMKLIK 182
Qy 131 KKAILKKKEKHIMSERNVLLKNVHPFLVGLHFSQFADKILYFVLDYINGGEIFYHLQR 190
Db 183 KEVIVAKDEVAHTLTE-NKVLONSRHPPLTALKYSFQTHDRLCFVMEYANGSELFHLSR 241
Qy 191 ERCLEPRARYAAEIASALGYLHS-LNIVYRDLKPNILLDSOGHIVLTDFGLCKENIE 249
Db 242 ERVSEDRARYGAEIVSALDYHSEKNVYRDVWMLCGLAVLYEMLYGLPPPYSRNTAEYDNI 309
Qy 250 HNSTTTCGTPPEYLAPVHLKQPDYDRTVDWMLCGLAVLYEMLYGLPPPYSRNTAEYDNI 309
Db 302 DGATMKTCGTPPEYLAPVLEDNDYGRAVDWMLGVLVYVWYEMMCGRLPYNDQHEKLELI 361
Qy 310 LNKPLQLKPNITNSARHLEGLLQKDKTKRL-GAKDDPMEIKSHVFFSLINWDLIINKKI 368
Db 362 LMEEIFRPTLGPPEAKSLSLGLLKKDPTQRLGGGSEDAKEIMQHRFFANIVMQDVEKKL 421
Qy 369 TTPPNPNVSGPNELRHFPETEEVPNSICKSPDSVLVTASVKEAARAFGLGF 424
Db 422 SPPPKQVTSYDTRFYDEEFTAQMITITPPDQDSMECVDS--ERRPHFPQFSYS 475

Search completed: August 11, 2003, 08:19:52
Job time : 46.5451 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:05:45 ; Search time 57.3433 Seconds
(without alignments)
722.816 Million cell updates/sec

Title: US-10-000-039a-2
Perfect score: 2270
Sequence: 1 MTKTEAAKGLTYSRMKGM.....KEAREAFGLGSFAPPTDSFL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2222	97.9	431	2 A48094	serum and glucocorticoid-regulated kinase - rat
2	1072	47.2	422	2 T22856	hypothetical prote
3	909	40.0	479	1 A59380	protein kinase (EC
4	905	39.9	462	1 T17287	protein kinase (EC
5	903	39.8	454	1 JC4345	protein kinase (EC
6	890	39.2	611	1 A55888	protein kinase (EC
7	865	38.1	480	1 S33364	protein kinase (EC
8	865	38.1	763	1 A40631	gag-akt polyprote
9	863	38.0	480	1 JC2437	protein kinase (EC
10	860.5	37.9	479	2 A38578	protein kinase (EC
11	850	37.4	480	1 A39360	protein kinase (EC
12	849	37.4	481	1 A46288	protein kinase (EC
13	847	37.3	480	1 S62117	protein kinase (EC
14	838	36.9	481	1 JC2438	protein kinase (EC
15	824	36.3	569	2 T50414	probable prolifera
16	823.5	36.3	680	2 S37955	protein kinase YPK
17	818.5	36.1	677	2 JS0178	protein kinase YKR
18	809.5	35.7	546	1 T43233	protein kinase (EC
19	808	35.6	481	2 J60377	p70 S6 kinase (EC
20	803	35.4	525	1 A41687	ribosomal protein
21	803	35.4	525	1 S12906	probable ribosomal
22	797	35.1	541	1 T43232	protein kinase (EC
23	795	35.0	634	1 B32392	protein kinase C (
24	794	35.0	525	1 TVRTK6	ribosomal protein
25	792.5	34.9	1016	1 A46079	protein kinase C (
26	791	34.8	672	1 KIHUCA	protein kinase C (
27	791	34.8	672	1 KIRTC	protein kinase C (
28	791	34.8	672	1 KINSCA	protein kinase C (
29	790	34.8	672	1 KIRBC	protein kinase C (

30 788.5 34.7 672 1 KIBOC protein kinase C (

31 780 34.4 547 2 T22856 hypothetical prote

32 780 34.4 671 1 KIBOC1 protein kinase C (

33 779 34.3 671 1 KIRBC1 protein kinase C (

34 777 34.2 707 1 A53530 protein kinase C (

35 776 34.2 671 1 KIRTC1 protein kinase C (

36 775.5 34.2 676 2 A17237 protein kinase C (

37 775.5 34.2 988 1 S35362 protein kinase C (

38 774.5 34.1 737 1 KIMSC protein kinase C (

39 773.5 34.1 587 2 A49509 protein kinase C (

40 772 34.0 696 2 S55694 protein kinase (EC

41 772 34.0 737 1 S28942 protein kinase C (

42 771.5 34.0 586 2 A53758 protein kinase C (

43 768.5 33.9 736 1 KIRBC protein kinase C (

44 767 33.8 528 1 T21523 protein kinase (EC

45 766 33.7 682 1 KIBOC protein kinase C (

ALIGNMENTS

RESULT 1

A48094

serum and glucocorticoid-regulated kinase - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 19-Dec-1997

C:Accession: A48094

R:Webster, M.K.; Goya, L.; Ge, Y.; Maiyar, A.C.; Firestone, G.L.

Mol. Cell. Biol. 13, 2031-2040, 1993

A:Title: Characterization of sgk, a novel member of the serine/threonine protein kinase

A:Reference number: A48094; MUID:93204949; PMID:8455596

A:Accession: A48094

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-431 <WEB>

A:Experimental source: Con8.hd6 mammary epithelial tumor cells

A:Note: sequence extracted from NCBI Backbone (NCBIN:127618, NCBIPI:127619)

C:Superfamily: Unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

C:Keywords: ATP

F:96-355/Domain: protein kinase homology <KIN>

F:104-112/Region: protein kinase ATP-binding motif

Query Match	97.94	Score	2222	DB 2	Length	431			
Best Local Similarity	97.08	Pred. No.	2.1e-94						
Matches	418	Conservative	8	Mismatches	5	Indels	0	Gaps	0
QY	1	MTVTEAAKGLTYSRMKGMVAIIAFMKQRRMGLNDFIQIANNYSACKHPEVQSTLKI	60						
DB	1	MTVKTEAARSTLTYSRMRGMVAIIAFMKQRRMGLNDFIQIANNYSACKHPEVQSTLKI	60						
QY	61	SOQPEELMNANPSPSPSQOINLGPSNPHAKPSDFHLKLVKIGKSGFKVLLARUKAE	120						
DB	61	SOQPEELMNANPSPSPSQOINLGPSNPHAKPSDFHLKLVKIGKSGFKVLLARUKAE	120						
QY	121	EYFVAVKVLQKAILKKKEKHIMSEKHNVLNKKVHPFLVGLHPSFOTADKLYFVLDYIN	180						
DB	121	EAFYAVKVLQKAILKKKEKHIMSEKHNVLNKKVHPFLVGLHFSFOTADKLYFVLDYIN	180						
QY	181	GGELFYHLQREKCFLEPRARYAAETASALGYLHSLNIVYRDLPENILDSQGHVILTD	240						
DB	181	GGELFYHLQREKCFLEPRARYAAETASALGYLHSLNIVYRDLPENILDSQGHVILTD	240						
QY	241	FGLCKENITEHNTSTTFCGTPYLAPEVLHQPVDRTVDMWCLGAVLYEMLYGLPPYSR	300						
DB	241	FGLCKENITEHNTSTTFCGTPYLAPEVLHQPVDRTVDMWCLGAVLYEMLYGLPPYSR	300						
QY	301	NTAEMYDNLNKPQLQKPNITNSARHLLLEGLLQDKRTKRLGAKDDFMEIKSHVFFSLINW	360						
DB	301	NTAEMYONILNKPQLQKPNITNSARHLLLEGLLQDKRTKRLGAKDDFMEIKSHVFFSLINW	360						
QY	361	DDLINKKITPPNPVNSGPNELRHDPDEFTPEPVPNSIGKSPDSVLTASVKEAEEAFLG	420						
DB	361	DDLINKKITPPNPVNSGPNLDLRHDPDEFTPEPVPSSIGKSPDSVLTASVKEAEEAFLG	420						

A:Residues: 1-462 <POU>
 A:Cross-references: EMBL:AL117525; GB:CA55977; NID:g5912043; PIDN:CA55977.1
 A:Experimental source: adult testis; clone DKF2p434N0250
 R:Nakatani, K.; Thompson, D.A.; Barthel, A.; Sakaue, H.; Liu, W.; Weigel, R.J.; Roth, R.
 J. Biol. Chem. 274, 21528-21532, 1999
 A:Title: Up-regulation of Akt3 in estrogen receptor-deficient breast cancers and androgen
 A:Reference number: A64199; PMID:10419456
 A:Contents: annotation
 A:Comment: This protein is increased in estrogen receptor-negative breast cancers and an
 C:Genetics:
 A:Gene: GDB:AKT3
 A:Cross-references: GDB:9954867
 A:Map position: 1q44-1q44
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
 A:Pathway: signal transduction pathways regulating various processes
 C:Superfamily: protein kinase akt; pleckstrin repeat homology; protein kinase homology
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein
 F:4-105/Domain: pleckstrin repeat homology <PLK>
 F:146-405/Domain: protein kinase homology <KIN>
 F:154-162/Region: protein kinase ATP-binding motif
 F:177/Active site: Lys #status predicted

Query Match 39.9%; Score 905; DB 1; Length 462;
 Best Local Similarity 47.6%; Pred. No. 1.3e-34;
 Matches 185; Conservative 61; Mismatches 115; Indels 28; Gaps 7;
 QY 28 MKQRMGLNDFIQK-----IANNYSACKHPE-----VQSILKISQPOPELMNANP 73
 DB 62 MKTERPKNTFTIIRCLQWTTVIERTFHVDTPEEREWEATEAQAVADRLQROEERMNKSP 121
 QY 74 SPPSPSQOINLG----PSSNPHAK---PSDFHLKLVKGSFGKVLARHKAEEVYAV 126
 DB 122 T-----SQIDNIGEEEMDASTTHHKRTMNDFDYLLKLGKGTGFKVILVREKASKYYAM 176
 QY 127 KVLQKKAILKKKEKHIMSERVLLKNVKKHPLVGLHFSFOTADKLYEVLQYINGGELFY 186
 DB 177 KILKKEVLIADKVAHTLTESRV-LKNTRHPFLTSLKYSFOTKORLCEVMEYVNGGELFF 235
 QY 187 HLQRCERCFLEPRARYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIVLTDFGLCKE 246
 DB 236 HLSRERVSSEDTREYGAIEVSALDYLSHGKIVYRDLPENILLDSQGHIVLTDFGLCKE 295
 QY 247 NLEHNSITSTFCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPYSRNTAEM 306
 DB 296 GITDAATMKTTCGTPPEYLAPEVLENDYGRAVDWMLGVMYEMMCGRLPFTYNDHEKLF 355
 QY 307 DNILNKLQLPKNTINSARHLEGLLOKDRKRL-GAKDDPMEIKSHVFFSLINWDDLIN 365
 DB 356 ELILMEDIKFPRTLSDDAKSLSGLLIKDPNKRLOGGPDPAKEIMRHSHFFSGVNNQDYYD 415
 QY 366 KKITPPPNPNVSGPNELRHDFPEEPEPV 394
 DB 416 KKLVPFPKPQVTSSETDTRYFDEEFTAQTI 444

RESULT 5
 JC4345
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
 N:Alternate names: protein kinase B gamma; RAC-PK-gamma; serine/threonine-specific prote
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
 C:Accession: JC4345
 R:Konishi, H.; Kuroda, S.; Tanaka, M.; Matsuzaki, H.; Ono, Y.; Kameyama, K.; Haga, T.; K
 Biochem. Biophys. Res. Commun. 216, 526-534, 1995
 A:Title: Molecular cloning and characterization of a new member of the RAC protein kinas
 e C subsequence and beta gamma subunits of G proteins.
 A:Reference number: JC4345; NID:96063640; PMID:7488143
 A:Accession: JC4345
 A:Molecule type: mRNA
 A:Residues: 1-454 <KON>
 A:Cross-references: DDBJ:D49836; NID:g1136777; PIDN:BAA08637.1; PID:g1401040
 A:Experimental source: brain

C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
 A:Pathway: signal transduction pathways regulating various processes
 C:Superfamily: protein kinase akt; pleckstrin repeat homology; protein kinase homology
 C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protei
 F:4-105/Domain: pleckstrin repeat homology <PLK>
 F:146-405/Domain: protein kinase homology <KIN>
 F:154-162/Region: protein kinase ATP-binding motif
 F:177/Active site: Lys #status predicted

Query Match 39.8%; Score 903; DB 1; Length 454;
 Best Local Similarity 47.6%; Pred. No. 1.6e-34;
 Matches 185; Conservative 61; Mismatches 115; Indels 28; Gaps 7;
 QY 28 MKQRMGLNDFIQK-----IANNYSACKHPE-----VQSILKISQPOPELMNANP 73
 DB 62 MKTERPKNTFTIIRCLQWTTVIERTFHVDTPEEREWEATEAQAVADRLQROEERMNKSP 121
 QY 74 SPPSPSQOINLG----PSSNPHAK---PSDFHLKLVKGSFGKVLARHKAEEVYAV 126
 DB 122 T-----SQIDNIGEEEMDASTTHHKRTMNDFDYLLKLGKGTGFKVILVREKASKYYAM 176
 QY 127 KVLQKKAILKKKEKHIMSERVLLKNVKKHPLVGLHFSFOTADKLYEVLQYINGGELFY 186
 DB 177 KILKKEVLIADKVAHTLTESRV-LKNTRHPFLTSLKYSFOTKORLCEVMEYVNGGELFF 235
 QY 187 HLQRCERCFLEPRARYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIVLTDFGLCKE 246
 DB 236 HLSRERVSSEDTREYGAIEVSALDYLSHGKIVYRDLPENILLDSQGHIVLTDFGLCKE 295
 QY 247 NLEHNSITSTFCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPYSRNTAEM 306
 DB 296 GITDAATMKTTCGTPPEYLAPEVLENDYGRAVDWMLGVMYEMMCGRLPFTYNDHEKLF 355
 QY 307 DNILNKLQLPKNTINSARHLEGLLOKDRKRL-GAKDDPMEIKSHVFFSLINWDDLIN 365
 DB 356 ELILMEDIKFPRTLSDDAKSLSGLLIKDPNKRLOGGPDPAKEIMRHSHFFSGVNNQDYYD 415
 QY 366 KKITPPPNPNVSGPNELRHDFPEEPEPV 394
 DB 416 KKLVPFPKPQVTSSETDTRYFDEEFTAQTI 444

RESULT 6
 A55888
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
 N:Alternate names: protein kinase B; RAC-PK; serine/threonine-specific protein kinase
 C:Species: Drosophila melanogaster
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 20-Apr-2001
 C:Accession: A55888
 R:Andjelkovic, M.; Jones, P.F.; Grossniklaus, U.; Cron, P.; Schlier, A.F.; Dick, M.; B
 J. Biol. Chem. 270, 4066-4075, 1995
 A:Title: Developmental regulation of expression and activity of multiple forms of the
 A:Reference number: A55888; NID:95181376; PMID:7876156
 A:Accession: A55888
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-611 <AND>
 A:Cross-references: GB:X83510
 C:Genetics:
 A:Gene: FlyBase:RacPK
 A:Cross-references: FlyBase:FBgn0013324
 A:Start codon: ACG
 A:Introns: 261/3; 327/3; 457/3; 535/3; 584/3
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threo
 A:Pathway: signal transduction pathways regulating various processes
 C:Superfamily: protein kinase akt; pleckstrin repeat homology; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/thre
 F:105-209/Domain: pleckstrin repeat homology <PLK>
 F:264-523/Domain: protein kinase homology <KIN>
 F:272-280/Region: protein kinase ATP-binding motif
 F:295/Active site: Lys #status predicted

QY 124 YAVKVLQKKAIIKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLDYINGGE 183
 DB 459 YAMKILKKEIVIAKDEVAHTLTE-NRVLQNSRHPFLTALKYSFQTHDRLCFVMEYANGGE 517
 QY 184 LFVHQREKCFLEPRARFYAAETASALGYLHS-LNIVYRDLPENILLDSQGHVILTDG 242
 DB 518 LFPHLSRERVFESDRARFYGAETVSALDYLSHSEKNVYRDLKLENLMDKDGHIKITTDFG 577
 QY 243 LCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFPYSRNT 302
 DB 578 LCKEGINDGATMKTFCGTPEYLAPEVLENDYGRADVWNLGVVYEMHMGCRLPFFNQDH 637
 QY 303 AEMYDNILANKPLQKNITNSARHLLEGLQKDKTKRL-GAKDDFMEIKSHVFFSLINWD 361
 DB 638 EKLFEILLMEIEIRPRTLGPPEAKSLLSGLLKKDKPTORLGGSEDAKEIMQHREFFANLVQ 697
 QY 362 DLINKKTIPTFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAAEAFLOF 421
 DB 698 DVEYKLLSPFPKQVTSSETDTRYFDEEFTAQMITITPPQDDDSMECVDS--ERRPHFPQF 755
 QY 422 SYA 424
 DB 756 SYS 758

RESULT 9

JC2437

protein kinase (EC 2.7.1.37) akt1 [validated] - rat
 N:Alternate names: protein kinase B alpha; RAC-PK-alpha; serine/threonine-specific prote
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 20-Apr-2001
 C:Accession: JC2437
 R:Konishi, H.; Shinomura, T.; Kuroda, S.; Ono, Y.; Kikkawa, U.
 Biochem. Biophys. Res. Commun. 205, 817-825, 1994
 A:Title: Molecular cloning of rat RAC protein kinase alpha and beta and their associatio
 A:Reference numbers: JC2437; MUID:95091823; PMID:7999118
 A:Accession: JC2437
 A:Molecule type: mRNA
 A:Residues: 1-480 <KON>
 A:Cross-references: DBJ:D30040; NID:G485402; PIDN:BAA06279.1; PID:G485403
 A:Experimental source: testis

C:Function:
 A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
 A:Pathway: signal transduction pathways regulating various processes
 C:Superfamily: protein kinase akt; pleckstrin repeat homology; protein kinase homology
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; proto-oncogene
 F:4-106/Domain: pleckstrin repeat homology <PLK>
 F:148-408/Domain: protein kinase homology <KIN>
 F:156-164/Region: protein kinase ATP-binding motif
 F:179/Active site: Lys #status predicted
 F:308/Binding site: phosphate (Thr) (covalent) (by phosphoinositide-dependent protein ki
 F:473/Binding site: phosphate (Ser) (covalent) (by autophosphorylation) #status predicte

Query Match 38.0%; Score 863; DB 1; Length 480;
 Best Local Similarity 43.8%; Pred. No. 1.1e-32;
 Matches 182; Conservative 75; Mismatches 137; Indels 22; Gaps 8;

QY 28 MKORMGLNDPIQK-----IANNYSACKHPE-----VOSTLKISQPOPELMNANP 73
 DB 63 MKTERPRNTTICLOWTIVERTFHVETPEEREEMTTAQTVAQDLKQREETMDFRS 122
 QY 74 SPPP--SPSQINIGPSSNPH-AKPSDFHLKLVCKSGFGKVLARHKAEEVFAVKVLQ 130
 DB 123 GSPSDNSGAEEVALAKPKHRTVWNEPEYLLKLGKGTGFKVILVKEKATGRYAMKTLK 182
 QY 131 KKAIIKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLDYINGGEFLYHQ 190
 DB 183 KEVIAKDEVAHTLTE-NRVLQNSRHPFLTALKYSFQTHDRLCFVMEYANGGEFLFLSR 241
 QY 191 ERCFLEPRARFYAAETASALGYLHS-LNIVYRDLPENILLDSQGHVILTDGFLCKENIE 249
 DB 242 ERVFSEDRARFYGAETVSALDYLSHSEKNVYRDLKLENLMDKDGHIKITTDFGLCKEGIK 301

QY 250 HNSTTTFCTGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFPYSRNTAEMYNIN 309
 DB 302 DCAATMLTKFCGTPEYLAPEVLENDYGRADVWNLGVVYEMHMGCRLPFFNQDHKEKFEI 361
 QY 310 LANKPLQKNITNSARHLLEGLQKDKTKRL-GAKDDFMEIKSHVFFSLINWDHINKKI 368
 DB 362 LMEIEIRPRTLGPPEAKSLLSGLLKKDKPTORLGGSEDAKEIMQHREFFANLVQJVEKKL 421
 QY 369 TTPFNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAAEAFLOF-SYA 424
 DB 422 SPPEKQVTSSETDTRYFDEEFTAQMITITPPQDDDSMECVDS--ERRPHFPQF-SYS 475

RESULT 10

A38578

protein kinase 2 (EC 2.7.1.1) - slime mold (Dictyostelium discoideum)
 C:Species: Dictyostelium discoideum
 C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 24-Sep-1999
 C:Accession: A38578
 R:Haribabu, B.; Dottin, R.P.
 Proc. Natl. Acad. Sci. U.S.A. 88, 1115-1119, 1991
 A:Title: Identification of a protein kinase multigene family of Dictyostelium discoid
 A:Reference number: A38578; MUID:91142122; PMID:1996312
 A:Accession: A38578
 A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-479 <HAR>

A:Cross-references: GB:M59744; NID:G167717; PIDN:AAA33186.1; PID:G167718
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/thre
 F:151-407/Domain: protein kinase homology <KIN>
 F:159-167/Region: protein kinase ATP-binding motif

Query Match 37.9%; Score 860.5; DB 2; Length 479;

Best Local Similarity 48.3%; Pred. No. 1.4e-32;

Matches 182; Conservative 57; Mismatches 111; Indels 27; Gaps 8;

QY 70 NANPPPPSPSQIN-----LGP-----SSNPHAKPSDFHLKLVCKSGFGKVLAKH 117
 DB 113 NEASSSPDSPNGSGNGDDEDEGPEEVIFSKNKSATKDDFELLNVICKSGFGKVMQVK 172
 QY 118 KAEEVFAVYVQKATLKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLD 177
 DB 173 KGEDTFAMKVLKRDALIAKQVNHVTKSEKTI-LQCISHPTFVNLHAFATKDKLYAVLD 231
 QY 178 YINGGELFYHQREKCELEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHV 237
 DB 232 FVNGGELFFHLKREGSEPRVKIYAAEIVSALDHLKQDIVYRDLPENILLDSQGHIC 291
 QY 238 LTDFGLCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFP 297
 DB 292 ITDFGLSKK-IETFDGTFCTGTPPEYLAPEVINGHGHGCAVDWMSLGTLLYFMITGIPPF 350
 QY 298 YSRNTAEMYNINLKNKPLQKNITNSARHLLEGLQKDKTKRLGAKDDFMEIKSHVFFSL 357
 DB 351 YSQNVNIMYQKILNGELKIPYISPEAKSLLEGTLTREVDRKLTGKG-GEYKUIPPEKN 409
 QY 358 INWDDLINKKITPPFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAA-- 415
 DB 410 IDWEKLORKEVEVHFEPKVKSGTDISQIDPVTQE-----RPMDSLVTESALGDAMGK 462

Query Match 37.9%; Score 860.5; DB 2; Length 479;

Best Local Similarity 48.3%; Pred. No. 1.4e-32;

Matches 182; Conservative 57; Mismatches 111; Indels 27; Gaps 8;

QY 70 NANPPPPSPSQIN-----LGP-----SSNPHAKPSDFHLKLVCKSGFGKVLAKH 117
 DB 113 NEASSSPDSPNGSGNGDDEDEGPEEVIFSKNKSATKDDFELLNVICKSGFGKVMQVK 172
 QY 118 KAEEVFAVYVQKATLKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLD 177
 DB 173 KGEDTFAMKVLKRDALIAKQVNHVTKSEKTI-LQCISHPTFVNLHAFATKDKLYAVLD 231
 QY 178 YINGGELFYHQREKCELEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHV 237
 DB 232 FVNGGELFFHLKREGSEPRVKIYAAEIVSALDHLKQDIVYRDLPENILLDSQGHIC 291
 QY 238 LTDFGLCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFP 297
 DB 292 ITDFGLSKK-IETFDGTFCTGTPPEYLAPEVINGHGHGCAVDWMSLGTLLYFMITGIPPF 350
 QY 298 YSRNTAEMYNINLKNKPLQKNITNSARHLLEGLQKDKTKRLGAKDDFMEIKSHVFFSL 357
 DB 351 YSQNVNIMYQKILNGELKIPYISPEAKSLLEGTLTREVDRKLTGKG-GEYKUIPPEKN 409
 QY 358 INWDDLINKKITPPFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAA-- 415
 DB 410 IDWEKLORKEVEVHFEPKVKSGTDISQIDPVTQE-----RPMDSLVTESALGDAMGK 462

Query Match 37.9%; Score 860.5; DB 2; Length 479;

Best Local Similarity 48.3%; Pred. No. 1.4e-32;

Matches 182; Conservative 57; Mismatches 111; Indels 27; Gaps 8;

QY 70 NANPPPPSPSQIN-----LGP-----SSNPHAKPSDFHLKLVCKSGFGKVLAKH 117
 DB 113 NEASSSPDSPNGSGNGDDEDEGPEEVIFSKNKSATKDDFELLNVICKSGFGKVMQVK 172
 QY 118 KAEEVFAVYVQKATLKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLD 177
 DB 173 KGEDTFAMKVLKRDALIAKQVNHVTKSEKTI-LQCISHPTFVNLHAFATKDKLYAVLD 231
 QY 178 YINGGELFYHQREKCELEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHV 237
 DB 232 FVNGGELFFHLKREGSEPRVKIYAAEIVSALDHLKQDIVYRDLPENILLDSQGHIC 291
 QY 238 LTDFGLCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFP 297
 DB 292 ITDFGLSKK-IETFDGTFCTGTPPEYLAPEVINGHGHGCAVDWMSLGTLLYFMITGIPPF 350
 QY 298 YSRNTAEMYNINLKNKPLQKNITNSARHLLEGLQKDKTKRLGAKDDFMEIKSHVFFSL 357
 DB 351 YSQNVNIMYQKILNGELKIPYISPEAKSLLEGTLTREVDRKLTGKG-GEYKUIPPEKN 409
 QY 358 INWDDLINKKITPPFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAA-- 415
 DB 410 IDWEKLORKEVEVHFEPKVKSGTDISQIDPVTQE-----RPMDSLVTESALGDAMGK 462

Query Match 37.9%; Score 860.5; DB 2; Length 479;

Best Local Similarity 48.3%; Pred. No. 1.4e-32;

Matches 182; Conservative 57; Mismatches 111; Indels 27; Gaps 8;

QY 70 NANPPPPSPSQIN-----LGP-----SSNPHAKPSDFHLKLVCKSGFGKVLAKH 117
 DB 113 NEASSSPDSPNGSGNGDDEDEGPEEVIFSKNKSATKDDFELLNVICKSGFGKVMQVK 172
 QY 118 KAEEVFAVYVQKATLKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLD 177
 DB 173 KGEDTFAMKVLKRDALIAKQVNHVTKSEKTI-LQCISHPTFVNLHAFATKDKLYAVLD 231
 QY 178 YINGGELFYHQREKCELEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHV 237
 DB 232 FVNGGELFFHLKREGSEPRVKIYAAEIVSALDHLKQDIVYRDLPENILLDSQGHIC 291
 QY 238 LTDFGLCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFP 297
 DB 292 ITDFGLSKK-IETFDGTFCTGTPPEYLAPEVINGHGHGCAVDWMSLGTLLYFMITGIPPF 350
 QY 298 YSRNTAEMYNINLKNKPLQKNITNSARHLLEGLQKDKTKRLGAKDDFMEIKSHVFFSL 357
 DB 351 YSQNVNIMYQKILNGELKIPYISPEAKSLLEGTLTREVDRKLTGKG-GEYKUIPPEKN 409
 QY 358 INWDDLINKKITPPFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAA-- 415
 DB 410 IDWEKLORKEVEVHFEPKVKSGTDISQIDPVTQE-----RPMDSLVTESALGDAMGK 462

Query Match 37.9%; Score 860.5; DB 2; Length 479;

Best Local Similarity 48.3%; Pred. No. 1.4e-32;

Matches 182; Conservative 57; Mismatches 111; Indels 27; Gaps 8;

QY 70 NANPPPPSPSQIN-----LGP-----SSNPHAKPSDFHLKLVCKSGFGKVLAKH 117
 DB 113 NEASSSPDSPNGSGNGDDEDEGPEEVIFSKNKSATKDDFELLNVICKSGFGKVMQVK 172
 QY 118 KAEEVFAVYVQKATLKKKEKHIMSERVLLKNVHPFLVLGHFSQTAADKLYFVLD 177
 DB 173 KGEDTFAMKVLKRDALIAKQVNHVTKSEKTI-LQCISHPTFVNLHAFATKDKLYAVLD 231
 QY 178 YINGGELFYHQREKCELEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHV 237
 DB 232 FVNGGELFFHLKREGSEPRVKIYAAEIVSALDHLKQDIVYRDLPENILLDSQGHIC 291
 QY 238 LTDFGLCKENIEHNSSTTFCGTPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPFP 297
 DB 292 ITDFGLSKK-IETFDGTFCTGTPPEYLAPEVINGHGHGCAVDWMSLGTLLYFMITGIPPF 350
 QY 298 YSRNTAEMYNINLKNKPLQKNITNSARHLLEGLQKDKTKRLGAKDDFMEIKSHVFFSL 357
 DB 351 YSQNVNIMYQKILNGELKIPYISPEAKSLLEGTLTREVDRKLTGKG-GEYKUIPPEKN 409
 QY 358 INWDDLINKKITPPFPNPNVSPNELRHFDPEETPEPNSIGKSPDSVLTASYKAA-- 415
 DB 410 IDWEKLORKEVEVHFEPKVKSGTDISQIDPVTQE-----RPMDSLVTESALGDAMGK 462

RESULT 11

A39360

protein kinase (EC 2.7.1.37) akt1 [validated] - human
 N:Alternate names: protein kinase B alpha; RAC-PK-alpha; serine/threonine-specific pr
 C:Species: Homo sapiens (man)
 C:Date: 20-Mar-1992 #sequence_revision 12-May-1994 #text_change 30-Sep-2001
 C:Accession: A39360; S36389; S18000; S20836

63 MKTERPENTFIIRCLQMTTIVIEKTHVETPEEREQWTTAIOQTVAQGLAKKQEEEMDFKS 121
74 SPPSPS----QQINLCGSSNPHAKP-----SDPFLKVGKSGFCKVLLIARIKAEVE 123
123 GSPSDNGSAEMEVSL-----AKPKHKVTMEFEYKLLGKGTGKVIILVKEKATGY 175
124 YAVVLQOKKATLKKKEEKHIMSEKRVLLKNVKHPFLVGLHFSPQADKLYFVLIDYINGKE 183
176 YAMKILKEVIAKDEVAHTLJE-NKVLQNSRHPFTALKYSPQTDRKLCFVMEYANGKE 234
184 LFYHLQPERCFLEPRAFYAAEIASALGYLHS-LNIVYRDLPKENTILISGCHIVLIDKGE 242
235 LFFHLSREVESEDRARFYGAIEVSALDYLHSEKNVYRDLKLENLMLKDKGHKILIDEG 294
243 LCKENIEHNSTTSTFCGTPYLAPEVLHKGPIYRTVDMMCLGAVLYEMLYGLPPFYSRNF 302
295 LCKEGIKDGAATMKTECGTPYLAPEVLEUDNDYGNADVWMLGVVYVMEYMGCKRLLPFYQNH 354
303 AEMVDNTLNKPLQLKPNITNSARHLLEGILLOKURTEL-GAKDDFMEIKSHWFFSLINWD 361
355 EKIFELIIMEEIRFRTLGVAEGLSLGLKDKPKOKRLOGGSEDAKIMQHRFFAGIIVWQ 414
362 DLINKKITTPPNNPVNSGFNELRHDPDFTTEEPNNSIGKSPDSVLTASVKAAAEAFGLGF 421
415 HVEKKLSPPKQVTSSTDTRYDEEFTAQMITTPDQDSDMECVDS--ERRPIHPQGF 472
422 SYA 424
473 SYS 475
RESULT 12
A46288
protein kinase (EC 2.7.1.37) akt2 - human
N:Alternate names: protein kinase B beta; RAC-PK-beta; serine/threonine-specific
C:Species: Homo sapiens (man)
C:Date: 22-Sep-1993 *sequence_revision 12-May-1994 *text_change 31-Mar-2001
C:Accession: A46288
R:Cheng, J.Q.; Godwin, A.K.; Bellacosa, A.; Taguchi, T.; Franke, T.F.; HamillCo
Proc. Natl. Acad. Sci. U.S.A. 89, 9267-9271, 1992
A:Title: AKT2, a putative oncogene encoding a member of a subfamily of protein
A:Reference number: A46288; MUID:93028445; PMID:1409633
A:Accession: A46288
A:Molecule type: mRNA
A:Residues: 1-481 <HE>
A:Cross-references: GB:M95936; NID:g178325; PIDN:AAA58364.1; PID:g178426
A:Note: sequence extracted from NCHI backbone (NCBIR:115859)
C:Comment: This protein is amplified in some pancreatic, ovarian, and other cancer
C:Genetics:
A:Gene: GDB:AKT2
A:Cross-references: GDB:135660; OMIM:164731
A:Map position: 19q13.2-19q13.2
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl
A:Pathway: signal transduction pathways regulating various processes
C:Superfamily: protein kinase akt; pleckstrin repeat homology; protein kinase
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific
F:4-106/Domain: pleckstrin repeat homology <PLK>
F:150-409/Domain: protein kinase homology <KIN>
F:158-166/Region: protein kinase ATP-binding motif
F:181/Active site: Lys #status predicted
Query Match 37.4%; Score 849; DB 1; Length 481;
Best Local Similarity 47.4%; Pred. No. 4.7e-32;
Matches 170; Conservative 59; Mismatches 114; Indels 16; Gaps 4
QY 39 IOKINNSYACKHPEVQSILKISQPOEPELANANPPSPPSQOINLGPSNPHAK--PS 96
103 IQWVANS-----LKQRPAGEDPMDYKCGSFSDSDSTTEMEVAVSKARAKVTMN 150
QY 97 DFHFLKVGKSGFKVLLARKHAAEEVYAVVLQKKAILKKKKEKHIMSERVLLKNVKH 156
151 DFDFLLKLGKGTGFGKVLIVREKATGRYVAMKILKEVIAKDEVAHTVTSRVLONTNRH 209

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:36:34 ; Search time 18.1364 Seconds
(without alignments)
462.401 Million cell updates/sec

Title: US-10-000-039a-3

Perfect score: 103
Sequence: 1 DPFTEPVPNSIGKSPDS 19

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US10000039/runat_01082003_085419_27535/app_query.fasta_1.661
-DB=Issued Patents.NA -OPMT=fastap -SUPPLX=rni -MINMATCH=0.1 -LOOPCL=0
-LCOEPT=0 -UNITS=bits -SPARS=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039@cgn_1_1_145@runat_01082003_085419_27535 -NCPU=6 -ICPU=3
-NO_MMWP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	2311	2 US-08-712-709-6	Sequence 6, Appli
2	103	100.0	2311	3 US-09-111-444-6	Sequence 6, Appli
3	103	100.0	2311	3 US-09-541-228-6	Sequence 6, Appli
4	103	100.0	2311	4 US-09-016-434-772	Sequence 772, App
5	103	100.0	2370	4 US-09-031-295-1	Sequence 1, Appli
6	51	49.5	2286	4 US-09-073-009-9	Sequence 9, Appli
7	51	49.5	4403765	3 US-09-103-840A-2	Sequence 2, Appli
8	51	49.5	4411529	3 US-09-103-840A-1	Sequence 1, Appli
9	50	48.5	910	4 US-09-482-273-88	Sequence 88, Appli
10	50	48.5	55827	4 US-09-813-133A-3	Sequence 3, Appli
11	46	44.7	2933	3 US-09-262-773-209	Sequence 209, App
12	46	44.7	23071	3 US-09-262-773-210	Sequence 210, App

13	46	44.7	4403765	3 US-09-103-840A-2	Sequence 2, Appli
14	46	44.7	4411529	3 US-09-103-840A-1	Sequence 1, Appli
15	45	43.7	1116	3 US-09-104-308-2	Sequence 2, Appli
16	45	43.7	1116	3 US-09-321-981-2	Sequence 2, Appli
17	45	43.7	1116	4 US-09-739-861A-2	Sequence 2, Appli
18	45	43.7	1116	4 US-09-795-583-2	Sequence 2, Appli
19	45	43.7	1158	4 US-09-328-352-246	Sequence 246, App
20	45	43.7	1253	1 US-08-332-747-2	Sequence 2, Appli
21	45	43.7	1716	3 US-09-321-981-4	Sequence 4, Appli
22	45	43.7	1716	4 US-09-739-861A-4	Sequence 4, Appli
23	45	43.7	1716	4 US-09-795-583-4	Sequence 4, Appli
24	44.5	43.2	1230	4 US-09-252-991A-1433	Sequence 1433, Ap
25	44.5	43.2	2025	4 US-09-252-991A-1536	Sequence 1536, Ap
26	44	42.7	246	4 US-09-193-881-3	Sequence 3, Appli
27	44	42.7	257	4 US-09-193-881-2	Sequence 2, Appli
28	44	42.7	586	4 US-09-221-0178-20	Sequence 20, Appli
29	44	42.7	892	2 US-08-460-529B-1	Sequence 1, Appli
30	44	42.7	906	3 US-08-831-132-1	Sequence 1, Appli
31	44	42.7	906	3 US-09-416-150-1	Sequence 1, Appli
32	44	42.7	1623	4 US-09-134-001C-1868	Sequence 1868, Ap
33	44	42.7	1756	4 US-09-193-881-11	Sequence 11, Appli
34	44	42.7	1760	4 US-09-193-881-10	Sequence 10, Appli
35	44	42.7	2351	4 US-09-177-419C-3	Sequence 3, Appli
36	43	41.7	1725	1 US-08-257-073-14	Sequence 14, Appli
37	43	41.7	2473	2 US-08-540-804-3	Sequence 3, Appli
38	43	41.7	2473	2 US-08-218-265-3	Sequence 3, Appli
39	43	41.7	2473	3 US-08-521-872-3	Sequence 3, Appli
40	43	41.7	2473	3 US-08-590-399-3	Sequence 3, Appli
41	43	41.7	2643	4 US-09-399-913-56	Sequence 56, Appli
42	43	41.7	2690	1 US-08-188-228-61	Sequence 61, Appli
43	43	41.7	2690	1 US-08-332-643-55	Sequence 55, Appli
44	43	41.7	2690	1 US-08-332-638-61	Sequence 61, Appli
45	43	41.7	2824	4 US-07-757-022B-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-712-709-6
; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6
Alignment Scores:
Pred. No.: 4.63e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-000-039A-3 (1-19) x US-08-712-709-6 (1-2311)
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 1165 GACCCCGAGTTTACCGAAGAGCCCTGTCCCACTCCATTCGCAAGTCCCTGCACAGC 1221
RESULT 2
US-09-111-444-6
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111.444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712.709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-09-111-444-6
Alignment Scores:
Pred. No.: 4.63e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-000-039A-3 (1-19) x US-08-712-709-6 (1-2311)
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 1165 GACCCCGAGTTTACCGAAGAGCCCTGTCCCACTCCATTCGCAAGTCCCTGCACAGC 1221
RESULT 2
US-09-111-444-6
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111.444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712.709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-09-111-444-6
Alignment Scores:
Pred. No.: 4.63e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-000-039A-3 (1-19) x US-08-712-709-6 (1-2311)
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 1165 GACCCCGAGTTTACCGAAGAGCCCTGTCCCACTCCATTCGCAAGTCCCTGCACAGC 1221
RESULT 4
US-09-016-434-772
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-10-000-039A-3 (1-19) x US-09-111-444-6 (1-2311)
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 1165 GACCCCGAGTTTACCGAAGAGCCCTGTCCCACTCCATTCGCAAGTCCCTGCACAGC 1221
RESULT 4
US-09-016-434-772
```

Sequence 772, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 772:
SEQUENCE CHARACTERISTICS:
LENGTH: 2311 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: MMLR2DT01
CLONE: 477245
US-09-016-434-772

Alignment Scores:
Pred. No.: 4.63e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-000-039a-3 (1-19) x US-09-016-434-772 (1-2311)
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
DB 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCAACTCCATTTGCGAAGTCCCTGACG 1221

RESULT 5
US-09-031-295-1
Sequence 1, Application US/09031295
Patent No. 6526181
GENERAL INFORMATION:
APPLICANT: LANG, Florian
APPLICANT: WALDEGGER, TUBINGEN
TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,295
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 197-08-173.8
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 058315/0123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2370 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 43..1335
US-09-031-295-1

Alignment Scores:
Pred. No.: 4.78e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-000-039a-3 (1-19) x US-09-031-295-1 (1-2370).
QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
DB 1198 GACCCCGAGTTTACCGAAGAGCGCTGTCCCAACTCCATTTGCGAAGTCCCTGACG 1254

RESULT 6
US-09-073-009-9/c
Sequence 9, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009

TUBERCULOSIS AND

QY 2 ProGluPheThrGluGluProValProAsnSerIleGlyLysSerPro 17
|||||
Db 3136791 CGGGCAGCAGCGAGGAGCTAGCCCTAACGCCGCTGCCAARATCC 3136838
|||||

RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R. M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.06e+06 Length: 4411529
Score: 46.00 Matches: 9
Percent Similarity: 62.50% Conservative: 1
Best Local Similarity: 56.25% Mismatches: 6
Query Match: 44.66% Indels: 0
DB: 3 Gaps: 0

US-10-000-039A-3 (1-19) x US-09-103-840A-1 (1-4411529)

QY 2 ProGluPheThrGluGluProValProAsnSerIleGlyLysSerPro 17
|||||
Db 3142622 CGGGCAGCAGCGAGGAGCTAGCCCTAACGCCGCTGCCAARATCC 3142669
|||||

RESULT 15
US-09-104-308-2/c
; Sequence 2, Application US/09104308
; Patent No. 6187577
; GENERAL INFORMATION:
; APPLICANT: Jones, Brian E.
; APPLICANT: Van Der Kleij, Wilhelmus A.H.
; APPLICANT: Van Solingen, Piet
; APPLICANT: Weyler, Walter
; TITLE OF INVENTION: No. 6187577ei Cellulase Producing Actinomycetes,
; TITLE OF INVENTION: Cellulase Produced Therefrom and Method of Producing Same
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,308
; FILING DATE: 24-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/974,042
; FILING DATE: 19-NOV-1997
; ATTORNEY/AGENT INFORMATION:

; NAME: Stone, Christopher L.
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC539
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7555
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-104-308-2
Alignment Scores:
Pred. No.: 132 Length: 1116
Score: 45.00 Matches: 9
Percent Similarity: 60.00% Conservative: 0
Best Local Similarity: 60.00% Mismatches: 6
Query Match: 43.69% Indels: 0
DB: 3 Gaps: 0
US-10-000-039A-3 (1-19) x US-09-104-308-2 (1-1116)
QY 3 GluPheThrGluGluProValProAsnSerIleGlyLysSerPro 17
|||||
Db 737 GAGTTCAGCGCAGACCGGTGCGCCCTCCACCGCTCGAAGCCC 693
|||||

Search completed: August 11, 2003, 12:07:09
Job time : 2223.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:23:34 ; Search time 70.8182 Seconds
(without alignments)
724.239 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103

Sequence: 1 DPETEEVPVNSIGKSPDS 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US10000039/runat_01082003.085418.27503/app_query.fasta.1.661
-DB=N_Geneseq_19Jun03 -ONET=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=prt -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039.ecgn.1.1.740@runat_01082003.085418.27503 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: N_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	1296	22	AAF44639
2	103	100.0	2281	20	AAV33647
3	103	100.0	2311	19	AAV23833
4	103	100.0	2311	25	ACA56174
5	103	100.0	2343	25	ABX74395
6	103	100.0	2346	20	AAV74190
7	103	100.0	2370	19	AAV48311
8	103	100.0	2370	24	ABK84749
9	103	100.0	2370	24	ABN97360
10	103	100.0	2370	24	ABL70006
11	100	97.1	404	25	ABX42261
12	100	97.1	433	25	ABX38963
13	95	92.2	431	24	ABI99409
14	95	92.2	2365	21	AAA46592
15	95	92.2	2432	22	AAAF4640
16	64	62.1	1333	22	AAH9535
17	64	62.1	1346	22	AAAF4641
18	64	62.1	1366	22	AAAF4987
19	64	62.1	1812	22	AAAF4737
20	64	62.1	1812	24	AAAF36141
21	64	62.1	1834	21	AAAF27856
22	64	62.1	2146	21	AAA27857
23	54	52.4	3953	22	AAAF37048
24	54	52.4	3953	25	ABX60036
25	51	49.5	2286	20	AAAF01152
26	51	49.5	2286	20	AAAF01186
27	51	49.5	2286	22	AAAF12464
28	51	49.5	2286	24	ABX63319
29	51	49.5	2965	22	AAAF48320
30	51	49.5	3872	24	ABAF1312
31	51	49.5	11809	22	ABAF21072
32	51	49.5	11809	22	AAAF07484
33	51	49.5	4403765	22	AAAF99683
34	51	49.5	4411529	22	AAAF99682
35	50	48.5	583	24	ABQ33516
36	50	48.5	583	24	ABQ33517
37	50	48.5	910	20	AAAF00677
38	50	48.5	910	21	AAAF98094
39	50	48.5	910	22	AAAF11707
40	50	48.5	910	24	ABK69803
41	50	48.5	910	25	ACC50366
42	50	48.5	910	25	ABZ71201
43	50	48.5	5915	24	ABL92310
44	50	48.5	5915	24	ABL49377
45	50	48.5	5915	24	AAAF22339

ALIGNMENTS

RESULT 1

AAF44639

ID AAF44639 standard; cDNA; 1296 bp.

XX

AC AAF44639;

XX

DT 27-MAR-2001 (first entry)

XX

DE Novel protein kinase cDNA, SEQ ID NO: 18.

XX

XX

KW Human; mouse; protein kinase; antiarthritic; osteopathic;

KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;

KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;

KW immune disorder; cardiovascular disease; neurodegenerative disease;

KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;

KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

XX

PA (INCY-) INCYTE PHARM INC.
XX Au-Young J, Guegler KJ, Hawkins PR;
XX WPI; 1998-207394/18.
DR P-PSDB; AAW54205.
XX New isolated human protein kinase(s) - used to develop products for
PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
PT asthma, multiple sclerosis or rheumatoid arthritis
XX Claim 25; Fig 3; 75pp; English.
XX This sequence encodes a human protein kinase (HPK) of the invention. The
CC HPK protein can be used to develop products for studying signalling
CC cascades in various cells and tissues; diagnosing disease and selecting
CC inhibitors or drugs with the potential to intervene in various disorders
CC or diseases in which altered kinase expression is implicated. The
CC products can be used to e.g. reverse memory loss such as due to
CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
CC lymphocytic leukaemias or lymphomas.
XX SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;
Alignment Scores:
Pred. No.: 9.67e-07 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0
US-10-000-039A-3 (1-19) x AAV23833 (1-2311)
Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCAACTCCATTGCGCAAGTCCCTGACAGC 1221
RESULT 4
ACA56174
ID ACA56174 standard; cDNA: 2311 BP.
XX ACA56174;
XX 06-JUN-2003 (first entry)
XX Human signalling pathway polynucleotide probe SEQ ID NO 772.
XX Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX Homo sapiens.
XX US6500938-B1.
XX 31-DEC-2002.
XX 30-JAN-1998; 98US-0016434.
XX 30-JAN-1998; 98US-0016434.
XX (INCY-) INCYTE GENOMICS INC.
XX Au-Young J, Seilhamer JJ;
XX WPI; 2003-352189/33.
XX Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides

XX Claim 1; SEQ ID NO 772; 65pp; English.
XX The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signalling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
XX SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;
Alignment Scores:
Pred. No.: 9.67e-07 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 25 Gaps: 0
US-10-000-039A-3 (1-19) x ACA56174 (1-2311)
Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCAACTCCATTGCGCAAGTCCCTGACAGC 1221
RESULT 5
ABX74395
ID ABX74395 standard; cDNA: 2343 BP.
XX ABX74395;
XX 21-MAR-2003 (first entry)
XX Human cDNA sequence #6 up-regulated in non-aggressive CC-RCC.
XX Human; microarray; solid surface; immobilised probe; CC-RCC;
KW differential expression profile; aggressive CC-RCC tumour type;
KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
KW gene expression profiling; tumour tissue; gene; ss.
XX Homo sapiens.
XX WO200279411-A2.
XX 10-OCT-2002.
XX 29-MAR-2002; 2002WO-US09576.
XX 29-MAR-2001; 2001US-279411P.
XX (VAND-) VAN ANDEL INST.
XX Haab B, Rhodes D, Teh BT, Takashi M;
XX WPI; 2003-040679/03.
XX

PT New microarray, comprising a matrix of cDNA probe from a set of probes
 PT immobilised to a solid surface in predetermined order, useful in the
 PT prognosis of patients with clear cell renal carcinoma -
 XX Claim 1: Page 79-80; 179pp; English.
 XX The present invention relates to a microarray comprising a matrix of
 CC at least one cDNA probe from a set of probes immobilised to a solid
 CC surface in a predetermined order, where a row of pixels corresponds
 CC to replicates of one distinct probe from the set. The probes are
 CC complementary to nucleic acid sequences that are expressed
 CC differentially in aggressive as compared to non-aggressive types of
 CC clear cell renal carcinoma (CC-RCC) and that hybridise to the probes
 CC under high stringency conditions. The microarray is useful for the
 CC prognosis of patients with CC-RCC, wherein aggressive and
 CC non-aggressive CC-RCC tumour types are characterised by differential
 CC expression profiles of genes that hybridise with one or more probes
 CC immobilised on the microarray. The arrays are useful for gene
 CC expression profiling of tumour and normal tissues. The present
 CC sequence represents a human cDNA sequence up-regulated in
 CC non-aggressive CC-RCC phenotypes.

SQ Sequence 2343 BP; 612 A; 517 C; 511 G; 703 T; 0 other;

Alignment Scores:
 Pred. No.: 9.83e-07 Length: 2343
 Score: 103.00 Matches: 19
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 25 Gaps: 0

US-10-000-039A-3 (1-19) x ABX74395 (1-2343)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysserProAspSer 19
 DB 1191 GACCCCGAGTTACCGAGAGCGCTGTCCCAACTCCATTCGCGAAGTCCCTGACAGC 1247

RESULT 6

AAV74190
 ID AAV74190 standard; DNA; 2346 BP.
 AC AAV74190;
 DT 15-MAR-1999 (first entry)
 XX Human sgk DNA.
 XX Sexum glucocorticoid regulated kinase; sgk; human; treatment; inhibitor;
 KW serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 KW chronic renal failure; inflammation; Alzheimers disease; wound; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 45..1340
 FT /*tag= a
 FT /product= "sgk"
 FT /trans_except= (pos:1185..1187, aa:Asp)
 XX EP887081-A2.
 XX 30-DEC-1998.
 XX 27-MAY-1998; 98EP-0304189.
 XX 27-JUN-1997; 97US-0051124.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Kumar JM;
 XX WPI; 1999-047627/05.

DR P-PSDB; AAW90139.

XX Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 PT disease - by administration of nucleic acids and antagonists which
 PT inhibit activity or expression of human serum glucocorticoid
 PT regulated kinase (sgk), a serine/threonine protein kinase
 XX Disclosure; Page 14-15; 17pp; English.
 XX This sequence encodes a novel human serum glucocorticoid regulated kinase
 CC (sgk) protein which is a member of the serine/threonine protein kinase
 CC family. This protein is used for the treatment of a subject having need
 CC to inhibit/antagonise activity or expression of human sgk polypeptide
 CC e.g. for the treatment of chronic renal failure, diabetic nephropathy,
 CC inflammation, Alzheimers disease and wounds.

SQ Sequence 2346 BP; 608 A; 520 C; 518 G; 700 T; 0 other;

Alignment Scores:
 Pred. No.: 9.85e-07 Length: 2346
 Score: 103.00 Matches: 19
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0

US-10-000-039A-3 (1-19) x AAV74190 (1-2346)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysserProAspSer 19
 DB 1200 GACCCCGAGTTACCGAGAGCGCTGTCCCAACTCCATTCGCGAAGTCCCTGACAGC 1256

RESULT 7

AAV48311
 ID AAV48311 standard; cDNA; 2370 BP.
 XX AAV48311;
 DT 16-NOV-1998 (first entry)
 XX Human cell-volume regulating kinase h-sgk.
 DE ss; human; cell-volume; kinase; h-sgk; diabetes mellitus;
 KW renal insufficiency; inflammation; Alzheimer's disease.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 43..1338
 FT /*tag= a
 FT /product= "Kinase h-sgk"
 XX EP861896-A2.
 XX 02-SEP-1998.
 XX 27-JAN-1998; 98EP-0101348.
 XX 28-FEB-1997; 97DE-1008173.
 XX (DADE-) DADE BEHRING MAKHURG GMBH.
 XX Lang F, Waldegger S;
 XX WPI; 1998-449109/39.
 XX P-PSDB; AAW77219.
 XX New nucleic acid encoding cell-volume regulating kinase h-sgk and
 PT related proteins - used for diagnosis and treatment of diseases
 PT involving changes in cell volume, e.g. renal insufficiency,
 PT inflammation, infections etc.
 XX Disclosure; Fig 1; 15pp; German.

XX The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.
 XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:
 Pred. No.: 9,96e-07 Length: 2370
 Score: 103.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0

US-10-000-039A-3 (1-19) x AAV48311 (1-2370)

Qy 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
 |||||
 Db 1198 GACCCGAGTTTACCGAAGAGCGCTGTCGCAACTCCATTGGCAAGTCCCTGCAGACG 1254

RESULT 8

ABK84749

ID ABK84749 standard; cDNA; 2370 BP.

XX AC ABK84749;

XX DT 14-AUG-2002 (first entry)

XX DE Human CDNA differentially expressed in granulocytic cells #1320.

XX KW Human; ss: granulocytic cell; DNA chip; bacterial infection;

XX KW viral infection; parasitic infection; protozoal infection;

XX KW fungal infection; sterile inflammatory disease; psoriasis;

XX KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

XX KW cardiac reperfusion injury; renal reperfusion injury; ARDS;

XX KW adult respiratory distress syndrome; inflammatory bowel disease;

XX KW Crohn's disease; ulcerative colitis; periodontal disease;

XX KW granulocyte activation; chronic inflammation; allergy.

XX OS Homo sapiens.

XX PN WO200228999-A2.

XX PD 11-APR-2002.

XX PF 03-OCT-2001; 2001WO-US30821.

XX PR 03-OCT-2000; 2000US-237189p.

XX PA (GENE-) GENE LOGIC INC.

XX PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX DR WPT; 2002-435328/46.

XX PT Detecting granulocyte activation by detecting differential expression

XX PT of genes associated with granulocyte activation, which serves as

XX PT diagnostic markers that is useful for monitoring disease states and

XX PT drug toxicity

XX PS Claim 1; SEQ ID No 1320; 114pp; English.

XX CC The invention relates to detecting (M1) granulocyte (GC) activation

XX CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by

XX CC DNA chip analysis as given in the specification, and comparing

CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicating GC of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease, by detecting the
 CC level of expression in a sample of the tissue of gene(s) from Gs, where
 CC the level of expression of the gene is indicative of inflammation;
 CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 CC an allergic response in a subject, exposure of a subject to a pathogen
 CC or sterile inflammatory disease, by contacting a tissue having
 CC inflammation with an agent that modulates the expression of gene(s)
 CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 CC modulating GCA; M3 is useful for screening an agent capable of modulating
 CC GCA preferably in an inflammation in a tissue; M4 is useful for
 CC detecting an inflammation (especially chronic) in a tissue, an allergic
 CC response in a subject, exposure of a subject to a pathogen or sterile
 CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 CC reperfusion injury, ARDS, adult respiratory distress syndrome,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC periodontal disease; also bacterial infection, viral infection, and M5 is
 CC parasitic infection, protozoal infection, fungal infection and M5 is
 CC useful for treating one of the above conditions. The present
 CC sequence represents a gene differentially expressed in granulocytes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:

Pred. No.: 9,96e-07 Length: 2370
 Score: 103.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-000-039A-3 (1-19) x ABK84749 (1-2370)

Qy 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
 |||||
 Db 1198 GACCCGAGTTTACCGAAGAGCGCTGTCGCAACTCCATTGGCAAGTCCCTGCAGACG 1254

RESULT 9

ABN97360

ID ABN97360 standard; DNA; 2370 BP.

XX AC ABN97360;

XX DT 13-AUG-2002 (first entry)

XX DE Gene #3858 used to diagnose liver cancer.

XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

XX KW metastatic liver tumour; cytostatic; expression profile; disease state;

XX KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX OS Homo sapiens.

XX PN WO200229103-A2.

XX PD 11-APR-2002.

XX PF 02-OCT-2001; 2001WO-US30589.

XX PR 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.
 PA Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 PI WPI; 2002-426119/45.
 DR
 XX Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample
 XX
 PS Claim 1; SEQ ID NO 3858; 298pp; English.
 XX
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
 Alignment Scores:
 Pred. No.: 9.96e-07 Length: 2370
 Score: 103.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-10-000-039A-3 (1-19) x ABN97360 (1-2370)
 Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
 DB 1198 GACCCCGAGTTACCGAGAGCGCTGTCCCAACATCCATTGGCAAGTCCCTTGACAGC 1254
 RESULT 10
 ABL70006
 ID ABL70006 standard: DNA; 2370 BP.
 XX
 AC ABL70006;
 XX
 DT 15-MAY-2002 (first entry)
 DE
 XX Pancreas cancer related gene sequence SEQ ID NO:8343.
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 PN WO2001194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-UST10838.
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 27-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 03-OCT-2000; 2000US-237609P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horriquan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 DR
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 8343; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
 Alignment Scores:

Pred. No.: 9.96e-07 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-10-000-039A-3 (1-19) x ABL70006 (1-2370)

QY 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 1198 GACCCCGAGTTACCGAAGAGCCGTGCCCCAACCTCCATTTGGCAAGTCCCTGGACAGC 1254

RESULT 11
ABX42261
ID ABX42261 standard; cDNA: 404 BP.
XX AC ABX42261;
XX
XX
XX 20-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #7426.
XX KW Bovine; ss: EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-0960352.
XX PR 12-JAN-1999; 99US-115707P.
XX PR 11-JAN-2000; 2000US-0480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI: 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2: SEQ ID No 7426; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived to a
CC from cattle, and the LMFD nucleic acid can specifically hybridize to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridization between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome

CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docid=20020137139.
XX
XX SQ Sequence 404 BP; 97 A; 108 C; 101 G; 98 T; 0 other;

Alignment Scores: 3.78e-07 Length: 404
Pred. No.: 100.00 Matches: 18
Score: 100.00% Conservative: 1
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 94.74% Indels: 0
Query Match: 97.05% Gaps: 0
DB: 25

US-10-000-039A-3 (1-19) x ABX42261 (1-404)

QY 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
DB 184 GATCTGAGTTCACTGAAGAGCCGTGCCCAACTCCATCCGCGCCGCGAGC 240

RESULT 12
ABX38963
ID ABX38963 standard; cDNA: 433 BP.
XX AC ABX38963;
XX
XX 20-FEB-2003 (first entry)
XX DT Bovine EST associated with lactation/muscle/fat deposition #4124.
XX DE Bovine EST associated with lactation/muscle/fat deposition #4124.
XX KW Bovine; ss: EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-0960352.
XX PR 12-JAN-1999; 99US-115707P.
XX PR 11-JAN-2000; 2000US-0480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI: 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and
PT analysis, cattle breeding, or for genetically improving cattle
XX
XX Claim 2: SEQ ID No 4128; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived to a
CC from cattle, and the LMFD nucleic acid can specifically hybridize to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49947, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3' end
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
CC complement or fragment) with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridization between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome

CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139.

SQ Sequence 433 BP; 107 A; 115 C; 105 G; 106 T; 0 other;

Alignment Scores:
 Pred. No.: 4,11e-07 Length: 433
 Score: 100.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 94.74% Mismatches: 0
 Query Match: 97.09% Indels: 0
 DB: 25 Gaps: 0

US-10-000-039A-3 (1-19) x ABX38963 (1-433)

QY 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19
 DB 184 GATCTGAGTTCTGAGAGCGGTCCTCCCACTCCATCGCGCGTCCCGGACAGC 240

RESULT 13

ABI99409
 ID ABI99409 standard; cDNA; 431 BP.

AC ABI99409;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:325.

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO2001188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

WP1: 2002-034733/04.

XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes.

XX Claim 2: Page 893; 2690pp; English.

PS The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular

CC genes (1) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (1). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in AB857020 to AB857374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.

SQ Sequence 431 BP; 75 A; 128 C; 115 G; 111 T; 2 other;

Alignment Scores:
 Pred. No.: 2,61e-06 Length: 431
 Score: 95.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 89.47% Mismatches: 0
 Query Match: 92.23% Indels: 0
 DB: 24 Gaps: 0

US-10-000-039A-3 (1-19) x ABI99409 (1-431)

QY 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19
 DB 15 GATCCCGAGTTTACCGAGAGCGCGTCCCGAGTCCATCGGAGTCCCTGACAGC 71

RESULT 14

AAA46592

ID AAA46592 standard; DNA; 2365 BP.

XX AC AAA46592;

XX 25-SEP-2000 (first entry)

DE DNA encoding a rat serum and glucocorticoid induced protein kinase.

XX Protein kinase; Pkh1; Pkh2; Ypk1; Ypk2; protein kinase B-alpha;

KW serum and glucocorticoid induced protein kinase; SGK; PKRalpha;

KW 3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
 KW thrush; cancer; diabetes; obesity; antifungal; Candida infection; ss.

XX Rattus sp.

XX Key Location/Qualifiers

FT CDS 1..1288

FT /tag- a

FT /product- "serum and glucocorticoid induced protein
 FT kinase (SGK)"

FT /transl_except- (pos: 1, aa: Thr)

FT /note- "the codon encoding Met at position 1 and
 FT Pro at position 318 are not given"

XX WO200036135-A2.

XX PD 22-JUN-2000.

XX PF 14-DEC-1999; 99WO-GB04228.

XX PR 14-DEC-1998; 98US-0112114.

XX (MEDI-) MEDICAL RES COUNCIL,

PA (REGC) UNIV CALIFORNIA.

XX Thorner JW, Alessi DR, Tortance PD, Casamayor A;

XX WPI; 2000-442391/38.

XX P-FSDB; AAY93530.

XX Screening method identifying compounds which modulate protein kinase
 PT activity for use in treating fungal infections and cancer

```
XX PS Disclosure; Fig 12; 155pp; English.
XX CC The present sequence encodes a rat serum and glucocorticoid induced
XX CC protein kinase (SGK). The specification describes a screening method
XX CC to identify a compound which modulates the activity of protein kinases
XX CC from different sources, using host yeast cells. The method is used to
XX CC identify a compound which modulates (inhibits) the activity of a
XX CC protein kinase. Pkhl or Pkh2 phosphorylate and activate Ypk1, Yrk2,
XX CC SGK or protein kinase B-alpha (PKBalpha). 3-phosphoinositide-dependent
XX CC protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
XX CC Yrk2 or SGK but not PKBalpha or p70S6 kinase. Compounds identified by
XX CC the methods are used to treat fungal infections e.g. thrush, and to
XX CC treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the
XX CC activation of PKB by Pkhl. Compounds which activate PKB or PDK1 can be
XX CC used in the treatment of diabetes or obesity, and compounds which
XX CC inhibit a fungal functional homologue of Pkhl (Pkh1 or Pkh2) or SGK
XX CC (Ypk1 or Yrk2) can be used as an antifungal agent to treat Candida
XX CC infections, e.g. thrush.
XX SQ Sequence 2365 BP; 591 A; 580 C; 518 G; 676 T; 0 other;

Alignment Scores:
Pred. No.: 1-93e-05 Length: 2365
Score: 95.00 Matches: 17
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 92.23% Indels: 0
DB: 21 Gaps: 0

US-10-000-039A-3 (1-19) x AAA46592 (1-2365)

Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1148 GATCCGAGTTTACTAGGAGCGCGTCCCGAGCTCCATCGCGGATCCCGTACAGC 1204

RESULT 15
AAF44640
ID AAF44640 standard; cDNA; 2432 BP.
AC AAF44640;
XX 27-MAR-2001 (first entry)
XX DE Novel protein kinase cDNA, SEQ ID NO: 19.
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
XX immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
XX dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
XX immune disorder; cardiovascular disease; neurodegenerative disease;
XX cancer; autoimmune disorder; stroke; inflammatory bowel disease;
XX inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX Mus musculus.
XX WO200073469-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US14842.
XX 28-MAY-1999; 99US-0136503.
XX (SUGR-) SUGEN INC.
XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX WPI: 2001-032161/04.
XX P-PSDB; AAB65614.
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
XX treating immune-related diseases and disorders, cardiovascular disease,
XX neurodegenerative diseases and/or cancers -
```

```
XX PS Disclosure; Fig 2; 310pp; English.
XX CC The present sequence encodes a novel protein kinase. The nucleic acids
XX CC and the protein kinases they encode may be used in the treatment and
XX CC diagnosis of diseases associated with inappropriate kinase expression
XX CC such as immune-related diseases and disorders, cardiovascular disease,
XX CC neurodegenerative diseases and/or cancers. The nucleic acids and
XX CC complementary sequences may also be used as DNA probes in diagnostic
XX CC assays. The kinase polypeptides may be used as antigens in the production
XX CC of antibodies of kinase expression and activity. Anti-kinase antibodies
XX CC and kinase antagonists may also be used to down regulate kinase
XX CC expression and activity. Diseases related to kinase expression and
XX CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
XX CC disorders, complications of organ transplantation, myocardial infarction,
XX CC immune disorders, cardiomyopathies, strokes, renal failure,
XX CC oxidative-stress related disorders, chronic inflammatory bowel disease,
XX CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
XX CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
XX CC reproductive disorders.
XX SQ Sequence 2432 BP; 586 A; 586 C; 560 G; 699 T; 1 other;

Alignment Scores:
Pred. No.: 2e-05 Length: 2432
Score: 95.00 Matches: 17
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 92.23% Indels: 0
DB: 22 Gaps: 0

US-10-000-039A-3 (1-19) x AAF44640 (1-2432)

Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1227 GATCCGAGTTTACCGAGGAGCGGTCCCGAGCTCCATCGCGGATCCCGTACAGC 1283

Search completed: August 11, 2003, 08:53:00
Job time : 78.8182 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:24:50 : Search time 532 Seconds
(without alignments)
1461.058 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103
Sequence: 1 DPFTPEPVNSIGKSPDS 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xl
-Q/cgn2.1/USPTO.spool/US1000039/runat_01082003_085418_27511/app_query.fasta_1.661
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT5=bits -START=1 -END=-1 -MATRIX=biom62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039 -CGN_1.1_5569 -runat_01082003_085418_27511 -NCPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pi:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_wi:*
15: gb_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_ror:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DM	ID	Description
1	103	100.0	1296	6	AX056374	Sequence
2	103	100.0	2281	6	AX017284	Sequence
3	103	100.0	2281	6	AX524988	Sequence
4	103	100.0	2281	6	BD134455	Human nuc
5	103	100.0	2311	6	AF151390	Sequence
6	103	100.0	2311	6	AF270209	Sequence
7	103	100.0	2311	6	BD021928	Human pro
8	103	100.0	2354	6	AX553549	Sequence
9	103	100.0	2364	9	AK098509	Homo sapi
10	103	100.0	2370	6	AR179441	Sequence
11	103	100.0	2370	6	AX002570	Sequence
12	103	100.0	2370	6	AX337834	Sequence
13	103	100.0	2370	6	AX411211	Sequence
14	103	100.0	2370	9	HSRNASPK	Y10032 H.sapiens m
15	103	100.0	2382	9	AF153609	Homo sapi
16	103	100.0	2382	9	BC001263	Sequence
17	103	100.0	3196	6	AX713538	Sequence
18	103	100.0	3196	9	AK055077	Homo sapi
19	103	100.0	5718	9	HSSGK	AK000512 Homo sapi
20	103	100.0	113673	9	AL135839	Human DNA
21	99	96.1	1417	5	AF057138	Xenopus l
22	97	94.2	2470	5	SAC223715	Sequence
23	97	94.2	3105	5	SAC223716	Squalus a
24	95	92.2	431	6	AX305574	Sequence
25	95	92.2	1659	10	BC002222	Mus muscu
26	95	92.2	1920	4	AF139639	Oryctolag
27	95	92.2	2423	10	BC005720	Mus muscu
28	95	92.2	2426	10	AF205855	Mus muscu
29	95	92.2	2429	10	AF139638	Mus muscu
30	95	92.2	2432	6	AX056375	Sequence
31	95	92.2	2435	10	RATSGPK	L01624 Rattus norv
32	95	92.2	142609	2	AC118095	Rattus no
33	95	92.2	208405	2	AC114405	Mus muscu
34	95	92.2	216056	2	AC103256	Rattus no
35	95	92.2	220350	2	AC107530	Rattus no
36	95	92.2	240716	2	AC112280	Rattus no
37	92	89.3	2339	5	BC052134	Danio rer
38	92	89.3	180682	2	BX323831	Danio rer
39	66	64.1	151	4	AF317416	Canis fam
40	65	63.1	183140	10	AC124486	Mus muscu
41	64	62.1	1346	6	AX056376	Sequence
42	64	62.1	1812	6	AX056472	Sequence
43	64	62.1	1812	6	AX399684	Sequence
44	64	62.1	1812	6	AF169034	Homo sapi
45	64	62.1	1815	9	BC014037	Homo sapi

ALIGNMENTS

RESULT 1

```

AX056374
LOCUS AX056374 1296 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 18 from Patent WO0073469.
ACCESSION AX056374
VERSION AX056374.1 GI:12229081
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS PLOWMAN,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 18 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source
Location/Qualifiers
1..1296
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 345 a 333 c 293 g 325 t
ORIGIN
Alignment Scores:
Pred. No.: 9.38e-09 Length: 1296
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 0 Gaps: 0

US-10-000-039A-3 (1-19) x AX056374 (1-1296)

Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1156 GACCCGAGTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTGACAGC 1212
|||||

RESULT 2
AX017284
LOCUS AX017284 2281 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 39 from Patent WO9947669.
ACCESSION AX017284
VERSION AX017284.1 GI:10042202
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS SCHMITT,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences from tissue of breast tumors
JOURNAL Patent: WO 9947669-A 39 23-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKI CHRISTIAN (DE)
FEATURES
source
Location/Qualifiers
1..2281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 601 a 498 c 494 g 588 t
ORIGIN
Alignment Scores:
Pred. No.: 1.78e-08 Length: 2281
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 0 Gaps: 0

US-10-000-039A-3 (1-19) x AX017284 (1-2281)

Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1156 GACCCGAGTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTGACAGC 1154
|||||

RESULT 3
AX0524988
LOCUS AX0524988 2281 bp DNA linear PAT 21-NOV-2002
DEFINITION Sequence 39 from Patent EP1236799.
ACCESSION AX0524988
VERSION AX0524988.1 GI:25170070
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Specht,T., Hinzmann,B., Schmitt,A., Pillarsky,C., Dahl,E. and
Rosenthal,A.
TITLE Human nucleic acid sequences derived from breast tumor tissue
JOURNAL Patent: EP 1236799-A 39 04-SEP-2002;
metagen Pharmaceuticals GmbH (DE)
FEATURES
source
Location/Qualifiers
1..2281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 601 a 498 c 494 g 588 t
ORIGIN
Alignment Scores:
Pred. No.: 1.78e-08 Length: 2281
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 0 Gaps: 0

US-10-000-039A-3 (1-19) x AX0524988 (1-2281)

Qy 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1098 GACCCGAGTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTGACAGC 1154
|||||

RESULT 4
BD134455
LOCUS BD134455 2281 bp DNA linear PAT 18-SEP-2002
DEFINITION Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION BD134455
VERSION BD134455.1 GI:23229400
KEYWORDS JP 2002506643-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2281)
AUTHORS Specht,T., Hinzmann,B., Armin,S., Pirarski,C., Edgar,D. and
Rosenthal,A.
TITLE Human nucleic acid sequence originating in mammary tumor tissue
JOURNAL Patent: JP 2002506643-A 33 05-MAR-2002;
METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT OS Homo sapiens (human)
PN JP 2002506643-A/33
PD 05-MAR-2002
PF 19-MAR-1999 JP 2000536852
PR 20-MAR-1998 DE 198 13 839.3
PI THOMAS SPECHT, BERND HINZMANN, SCHMITT ARMIN, CHRISTIAN PILARSKI,
PI DUHL EDGAR,
PI ANDRE ROSENTHAL
PI C12N15/09, A61K31/711, A61K38/00, A61P35/00, C07K14/47,
PC C07K16/18,
PC C12N1/19, C12N5/10, C12N15/00, A61K37/02, C12N5/00 CC Human
nucleic acid sequence originating in mammary tumor CC

```

tissue
FH Key Location/Qualifiers
FT source 1..2281
FT /organism="Homo sapiens (human)"

FEATURES
source Location/Qualifiers
1..2281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 601 a 498 c 494 g 688 t
ORIGIN

Alignment Scores:
Pred. No.: 1 78e-08 Length: 2281
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x BD134455 (1-2281)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1098 GACCCCGAGTTTACCGAAGAGCGCTGTCCCACTCCATTGGCAAGTCCCTGACAGC 1154

RESULT 5
AR151390
LOCUS AR151390 2311 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 6 from patent US 6232077.
ACCESSION AR151390
VERSION AR151390.1 GI:15117440
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Au-Young, J., Guegler, K.J. and Hawkins, P.R.
TITLE Human protein kinases
JOURNAL Patent: US 6232077-A 6 15-MAY-2001;
FEATURES
source Location/Qualifiers
1..2311
/organism="unknown"

BASE COUNT 604 a 508 c 506 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1 81e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x AR151390 (1-2311)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCACTCCATTGGCAAGTCCCTGACAGC 1221

RESULT 6
AR270209
LOCUS AR270209 2311 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 772 from patent US 6500938.
ACCESSION AR270209
VERSION AR270209.1 GI:29701443
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Au-Young, J. and Seilhamer, J.J.
TITLE Composition for the detection of signaling pathway gene expression

JOURNAL Patent: US 6500938-A 772 31-DEC-2002;
FEATURES
source Location/Qualifiers
1..2311
/organism="unknown"

BASE COUNT 604 a 508 c 506 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1 81e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x AR270209 (1-2311)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCACTCCATTGGCAAGTCCCTGACAGC 1221

RESULT 7
BD021928
LOCUS BD021928 2311 bp DNA linear PAT 27-AUG-2002
DEFINITION Human protein kinases.
ACCESSION BD021928
VERSION BD021928.1 GI:22563151
KEYWORDS JP 2001506843-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2311)
AUTHORS Young, J.O., Gegura, K.J. and Horikawa, P.R.
TITLE Human protein kinases
JOURNAL Patent: JP 2001506843-A 3 29-MAY-2001;
INSITE PHARMACEUTICALS INC
COMMENT
PN JP 2001506843-A/3
PD 29-MAY-2001
PF 10-SEP-1997 JP 1998513776
PI 12-SEP-1996 US 08/712709
PI JANICE O YOUNG, KARL J GEGURA, PHILLIP R HORIKAWA
C12N15/00, A61K38/45, A61K39/395, A61K45/00, A61K48/00, PC
A61P11/06,
PC A61P19/02, A61P25/28, A61P29/00, A61P35/00, A61P43/00, C07K16/40,
PC C12N1/15, C12N1/21, C12N5/10, C12N9/12, C12O1/48, C12N15/00, C12N5/
PC 00, A61K37/52
CC

FEATURES
source Location/Qualifiers
1..2311
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32844"

BASE COUNT 604 a 507 c 507 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1 81e-08 Length: 2311
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x BD021928 (1-2311)

QY 1 AspProGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1165 GACCCCGAGTTTACCGAAGAGCGCTGTCCCACTCCATTGGCAAGTCCCTGACAGC 1221

RESULT 8
AX553549

LOCUS AX553549 2354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 1 from Patent WO02074987.
ACCESSION AX553549
VERSION AX553549.1 GI:25897549
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE
AUTHORS Lang, F., Busjahn, A. and Luft, F.C.
TITLE Quantitative diagnostic analysis of hypertension
JOURNAL Patent: WO 02074987-A 1 26-SEP-2002;
Lang, Florian (DE)
FEATURES
source 1..2354
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
43..1338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD58123.1"
/db_xref="GI:25897550"
translation="MTVTEAAKGLTYSRMRGMVAIIIAFMKORRMGLNDFIOKIAN
NSYKGSFQKVLARHAAEEFYAVKLVKKKAEKHKHMSRNVLNKNVHPFLV
VIGKSGFQKVLARHAAEEFYAVKLVKKKAEKHKHMSRNVLNKNVHPFLV
GLHFSQFADKLYFVLDINGELFYHLORERCFLEPRARFYAAEIASALCYLHSLNI
VYRDLKPNILDSQCHIVLDFGLCKENIEHNSSTFTCGPEYLAPEVLHQPYPOR
TVDMKGLAVGLYEMLYKPFYSRNTAEYDNLKPLQKPNITNSARLLGGLGK
DFTKLGAKDDFMELKSHVFSLNWDLLNKITPPENPNVSGPNDLRHFDPEFTEE
PVPNSIGSPDSVLYTASVKRAEAEFLGFSVAPPTDSFL"
762
/note="1. SNP (C in T), stumme Mutation, d.h. beide
Versionen des SNPs resultieren in der AMINO ACID Asp in
der AMINO ACID-Position 240"
variation 613 a 518 c 513 g 710 t
BASE COUNT 613 a 518 c 513 g 710 t
ORIGIN
Alignment Scores:
Pred. No.: 1,85e-08 Length: 2354
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039a-3 (1-19) x AX553549 (1-2354)
Qy 1 AspproGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1198 GACCCCGAGTTTACCGAGAGGCGCTGTCCCAACTCCATTGGCAAGTCCCTGCACAGC 1254
RESULT 9
LOCUS AX553549 2364 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FL25643 f1s, clone STM07148, highly similar to
SERINE/THREONINE-PROTEIN KINASE SGK (EC 2.7.1.1).
ACCESSION AK098509
VERSION AK098509.1 GI:21758535
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,
Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S.,
Morinaga, M., Kawamura, M., Sugiyama, T., Irie, K., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isodai, T. and
Sugano, S.

TITLE Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2364)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.
FEATURES
source 1..2364
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM07148"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/note="cloning vector: pHE185FL3"
BASE COUNT 615 a 524 c 517 g 708 t
ORIGIN
Alignment Scores:
Pred. No.: 1,86e-08 Length: 2364
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-000-039a-3 (1-19) x AK098509 (1-2364)
Qy 1 AspproGluPheThrGluGluProValProAsnSerIleGlyLysSerProAspSer 19
Db 1213 GACCCCGAGTTTACCGAGAGGCGCTGTCCCAACTCCATTGGCAAGTCCCTGCACAGC 1269
RESULT 10
LOCUS AR179441 2370 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6326181.
ACCESSION AR179441
VERSION AR179441.1 GI:20220996
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2370)
AUTHORS Lang, F. and Waldegger, S.
TITLE Cell volume-regulated human kinase h-sgk
JOURNAL Patent: US 6326181-A 1 04-DEC-2001;
FEATURES Location/Qualifiers
source 1..2370
/organism="unknown"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 1,86e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039a-3 (1-19) x AR179441 (1-2370)

QY 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTCGACAGC 1254

RESULT 11
AX002570
LOCUS AX002570 2370 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 1 from Patent EP0861896.
ACCESSION AX002570
VERSION AX002570.1 GI:7242111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2370)
AUTHORS Lang, F.P. and Waldegger, S.D.
TITLE Cell volume regulated human kinase h-sgk
JOURNAL Patent: EP 0861896-A 1 02-SEP-1998;
DADE BEHRING MARBURG GMBH (DE)
FEATURES
source Location/Qualifiers
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/strain="CAUCASIAN"
/db_xref="taxon:9606"
/clone="HSRNAS7PK"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 1.86e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x AX002570 (1-2370)

QY 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTCGACAGC 1254

RESULT 12
AX337834
LOCUS AX337834 2370 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 8343 from Patent WO0194629.
ACCESSION AX337834
VERSION AX337834.1 GI:18128553
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL gene sets
Patent: WO 0194629-A 8343 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source Location/Qualifiers
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 1.86e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x AX337834 (1-2370)

QY 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTCGACAGC 1254

RESULT 13
AX411211
LOCUS AX411211 2370 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3858 from Patent WO0229103.
ACCESSION AX411211
VERSION AX411211.1 GI:21443916
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3858 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
source Location/Qualifiers
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. Y10032"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 1.86e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-3 (1-19) x AX411211 (1-2370)

QY 1 AspProGluPheThrGluProValProAsnSerIleGlyLysSerProAspSer 19
|||||
Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTGGCAAGTCCCTCGACAGC 1254

RESULT 14
HSRNAS7PK
LOCUS HSRNAS7PK 2370 bp mRNA linear PRI 16-MAY-1997
DEFINITION H.sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION Y10032
VERSION Y10032.1 GI:1834510
KEYWORDS serine/threonine protein kinase; sgk gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Waldegger, S., Barth, P., Raber, G. and Lang, F.
TITLE Cloning and characterization of a putative human serine/threonine
protein kinase transcriptionally modified during anisotonic and
isotonic alterations of cell volume
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (9), 4440-4445 (1997)
MEDLINE 97272242
PUBMED 9114008
REFERENCE 2 (bases 1 to 2370)
AUTHORS Waldegger, S.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1996) S. Waldegger, University Of Tuebingen,

Physiology 1, Gmelinstr. 5, D-72076 Tuebingen, FRG

FEATURES

Location/Qualifiers
1. 2370

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="adult"

1. 2370

/gene="sgk"

43. 1338

/gene="sgk"

/codon_start=1

/product="serine/threonine protein kinase"

/protein_id="CAA71138.1"

/db_xref="GI:1834511"

/translation="MTVKTEAAKGTLYSRMRGVAIIAFMKORHMGCLNDFIOKIAN

NSYACKHPEVSILKISQPELMNANPSPSPSQIINLGPSNPHAKPSDFHFLK
VIGKSGFKVLLARHAEVYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFLV
GLHFSQADKLYFVLDYINGGELFYHLQRCERCFLEPRAREYAAEIASALGYLHSLNI
VYRDLKPENILDSOGHIVLTDGCLCKENIEHNSTTSTFCGTPEYLAPEVLHKOPYDR
TVDMWCLGAVLYEMLYGLPPFYSRNTAEYDNLNKLPLQKPNITNSARHLLEGLOK
DTRRLGAKDDFMEIKSHVFFSLINWDLINKITPPFNPNVSGPNELRHPDPETEE
PVPNSIGKSPDSVLVTASVKEAAEAFLGFSYAPPTDSFL"

BASE COUNT 636 a 517 c 513 g 704 t

ORIGIN

Alignment Scores:
Pred. No.: 1.86e-08 Length: 2370
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-000-039a-3 (1-19) x HSRNASTPK (1-2370)

Qy 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19

Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTCGCAAGTCCCTGCAGAC 1254

RESULT 15

AF153609

LOCUS

AF153609 Homo sapiens serine/threonine protein kinase sgk mRNA, complete

DEFINITION

AF153609

VERSION

AF153609.1 GI:5231142

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 2382)

Kim.M.K., Kim.Y.H., Seo.J.M., Lee.H.M., Chung.H.J., Sohn.M.Y.,

Hwang.S.Y., Im.S.U., Jung.E.J. and Kim.J.C.

A catalogue of genes in the human dermal papilla cells as

identified by expressed sequence tags

Unpublished

2 (bases 1 to 2382)

Kim.M.K., Kim.Y.H., Suh.J.M., Lee.H.M., Chung.H.J., Sohn.M.Y.,

Hwang.S.Y., Im.S.U., Jung.E.J. and Kim.J.C.

Direct Submission

Submitted (24-MAY-1999) Immunology, Kyungpook National University,

School of Medicine, 101 Dongin Dong, Jung Gu, Taegu, Taegu 700-422,

South Korea

Location/Qualifiers

1. 2382

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="primary culture"

/cell_type="hair dermal papilla"

FEATURES

source

CDS

43. 1338

/codon_start=1

/product="serine/threonine protein kinase sgk"

/protein_id="AAD41091.1"

/db_xref="GI:5231143"

/translation="MTVKTEAAKGTLYSRMRGVAIIAFMKORHMGCLNDFIOKIAN

NSYACKHPEVSILKISQPELMNANPSPSPSQIINLGPSNPHAKPSDFHFLK
VIGKSGFKVLLARHAEVYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFLV
GLHFSQADKLYFVLDYINGGELFYHLQRCERCFLEPRAREYAAEIASALGYLHSLNI
VYRDLKPENILDSOGHIVLTDGCLCKENIEHNSTTSTFCGTPEYLAPEVLHKOPYDR
TVDMWCLGAVLYEMLYGLPPFYSRNTAEYDNLNKLPLQKPNITNSARHLLEGLOK
DTRRLGAKDDFMEIKSHVFFSLINWDLINKITPPFNPNVSGPNELRHPDPETEE
PVPNSIGKSPDSVLVTASVKEAAEAFLGFSYAPPTDSFL"

BASE COUNT 640 a 519 c 513 g 710 t

ORIGIN

Alignment Scores:
Pred. No.: 1.87e-08 Length: 2382
Score: 103.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-000-039a-3 (1-19) x AF153609 (1-2382)

Qy 1 AspProGluPheThrGluCluProValProAsnSerIleGlyLysSerProAspSer 19

Db 1198 GACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCCATTCGCAAGTCCCTGCAGAC 1254

Search completed: August 11, 2003, 10:05:04

Job time : 535 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:19:59 : Search time 0.896996 Seconds
(without alignments)
424.584 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103
Sequence: 1 DPETPEVPNSICKSPDS 19

Scoring Table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 98705 seqs, 20044758 residues

Total number of hits satisfying chosen parameters: 98705

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	100.0	19	US-10-000-039A-3	Sequence 3, Appli
2	103	100.0	431	PCT-US03-10142-2	Sequence 2, Appli
3	103	100.0	431	PCT-US03-10142-4	Sequence 4, Appli
4	103	100.0	431	US-10-000-039A-2	Sequence 2, Appli
5	64	62.1	367	US-10-291-172-256	Sequence 256, App
6	46	44.7	414	US-10-603-113-15403	Sequence 15403, A
7	45	43.7	1706	US-10-293-244-1932	Sequence 1932, Ap
8	44	42.7	468	US-10-273-573-5843	Sequence 5843, Ap
9	44	42.7	496	US-10-286-897-1979	Sequence 1979, Ap
10	44	42.7	496	US-10-286-897-5551	Sequence 5551, Ap
11	44	42.7	1122	US-10-408-765A-1005	Sequence 1005, Ap
12	44	42.7	1080	US-10-307-817-138	Sequence 138, App
13	44	42.7	4624	US-10-408-765A-2991	Sequence 2991, Ap
14	43	41.7	687	US-10-326-956-2353	Sequence 2353, Ap
15	43	41.7	713	US-10-408-765A-394	Sequence 394, App
16	43	41.7	1415	US-10-273-573-9302	Sequence 9302, A
17	42	40.8	84	US-10-603-113-25178	Sequence 25178, A
18	42	40.8	228	US-10-613-520-1597	Sequence 1597, Ap
19	42	40.8	464	US-10-394-322A-13	Sequence 13, Appli
20	41.5	40.3	4618	US-10-286-897-2188	Sequence 2188, Ap
21	41	39.8	351	PCT-US03-21510-80	Sequence 80, Appli
22	41	39.8	351	US-10-293-244-1863	Sequence 1863, Ap
23	41	39.8	3859	US-10-408-765A-354	Sequence 354, App
24	40.5	39.3	324	US-10-603-113-15950	Sequence 15950, A
25	40.5	39.3	972	US-10-408-765A-2934	Sequence 2934, Ap
26	40.5	39.3	1190	US-10-326-956-3263	Sequence 3263, Ap

ALIGNMENTS

RESULT 1

US-10-000-039A-3

; Sequence 3, Application US/10000039A

; GENERAL INFORMATION:

APPLICANT: LANG, Florian

TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: FOLEY & IARDNER

STREET: 3000 K Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION NUMBER: US/10/000,039A

FILING DATE: 04-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/031,295

FILING DATE: 26-FEB-1998

APPLICATION NUMBER: DE 197-08-173.8

FILING DATE: 28-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sandercock, Colin G.

REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 058315/0123

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-000-039A-3

Query Match

100.0%; Score 103; DB 6; Length 19;

Best Local Similarity 100.0%; Pred. No. 4.3e-10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 1062, Ap
Sequence 1063, Ap
Sequence 1366, Ap
Sequence 14349, A
Sequence 6287, Ap
Sequence 2715, Ap
Sequence 974, App
Sequence 102, App
Sequence 4312, Ap
Sequence 2238, Ap
Sequence 972, App
Sequence 10231, A
Sequence 10331, A
Sequence 464, App
Sequence 3225, Ap
Sequence 1237, Ap
Sequence 1813, Ap
Sequence 3259, Ap
Sequence 5760, Ap

QY 1 DPEETEPVNSIGKSPDS 19
|||||
DB 1 DPEETEPVNSIGKSPDS 19

RESULT 2
PCT-US03-10142-2
; Sequence 2, Application PC/TUS0310142
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: PCT/US03/10142
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-10142-2

Query Match 100.0%; Score 103; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEPVNSIGKSPDS 19
|||||
DB 386 DPEETEPVNSIGKSPDS 404

RESULT 3
PCT-US03-10142-4
; Sequence 4, Application PC/TUS0310142
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-573C
; CURRENT APPLICATION NUMBER: PCT/US03/10142
; PRIOR FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/370349
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/384543
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/370969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/403748
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/372019
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/374379

; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: 09/779679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 10/055877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262892
; PRIOR FILING DATE: 2001-01-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-10142-4

Query Match 100.0%; Score 103; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEPVNSIGKSPDS 19
|||||
DB 386 DPEETEPVNSIGKSPDS 404

RESULT 4
US-10-000-039A-2
; Sequence 2, Application US/10000039A
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SOK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-IMS/MS-IVS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/000,039A
; FILING DATE: 04-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-000-039A-2

Query Match 100.0%; Score 103; DB 6; Length 431;

2525 20004 041111814245
33,18, 1103, 110, 122

Matches 9: Conservative 4: Mismatches 4: Indels 6: Gaps 1:

QY 1 DPEFTE-----EVPNSICKSP 17
|||||
DB 437 DPWTKALPPFPASPLPGFLGASP 459

RESULT 9

US-10-286-897-1979
; Sequence 1979, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 1979
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-1979

Query Match 42.7%; Score 44; DB 6; Length 496;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 DPEFTEPVPNSICKSP 18
|||||
DB 450 DTAFTEPVPNSICKSP 467

RESULT 10

US-10-286-897-5551
; Sequence 5551, Application US/10286897
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/10/286,897
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US/09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US/09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US/09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US/09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US/09/653,450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US/09/662,191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US/09/693,036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US/09/727,344

; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7143
; SOFTWARE: pt_FL_genes_b Versions 1.0
; SEQ ID NO 5551
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-897-5551

Query Match 42.7%; Score 44; DB 6; Length 496;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 DPEFTEPVPNSICKSP 18
|||||
DB 450 DTAFTEPVPNSICKSP 467

RESULT 11

US-10-408-765A-1005
; Sequence 1005, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary W.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1005
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1005

Query Match 42.7%; Score 44; DB 6; Length 1122;
Best Local Similarity 55.6%; Pred. No. 50;
Matches 10; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 DPEFTEPVPNSICKSP 18
|||||
DB 475 NPE--EELPNAYDKSLD 490

RESULT 12

US-10-307-817-138
; Sequence 138, Application US/10307817
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 138
; LENGTH: 4080
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-138

Query Match 42.7%; Score 44; DB 6; Length 4080;
Best Local Similarity 55.6%; Pred. No. 2,1e+02;
Matches 10; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 DPEFTEPVPNSICKSP 18

DB 3936 NPE--EEPLPNAYDKSLD 3951
 :|| ||:|: || |

RESULT 13

US-10-408-765A-2991
 ; Sequence 2991, Application US/10408765A
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Wainock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408.765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2991
 ; LENGTH: 4624
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-2991

Query Match 42.7%; Score 44; DB 6; Length 4624;
 Best Local Similarity 55.6%; Pred. No. 2.4e+02;
 Matches 10; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 1 DPEETPEVPVNSIGKSPD 18
 :|| ||:|: || |
 DB 3977 NPE--EEPLPNAYDKSLD 3992

RESULT 14

US-10-326-956-2353
 ; Sequence 2353, Application US/10326956
 ; GENERAL INFORMATION:
 ; APPLICANT: Bauer et al.
 ; TITLE OF INVENTION: Protein Complexes and Methods for their Use
 ; FILE REFERENCE: 220615
 ; CURRENT APPLICATION NUMBER: US/10/326.956
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: EP 01 130 253.6
 ; PRIOR FILING DATE: 2001-12-20
 ; NUMBER OF SEQ ID NOS: 3282
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2353
 ; LENGTH: 687
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-10-326-956-2353

Query Match 41.7%; Score 43; DB 6; Length 687;
 Best Local Similarity 53.3%; Pred. No. 41;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 TEETPEVPVNSIGKSPDS 19
 :|||| | | | |
 DB 43 SDEPVFESAGKADTS 57

RESULT 15

US-10-408-765A-394
 ; Sequence 394, Application US/10408765A
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Eoin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Wainock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408.765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 394
 ; LENGTH: 713
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-394

Query Match 41.7%; Score 43; DB 6; Length 713;
 Best Local Similarity 44.4%; Pred. No. 43;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PEETPEVPVNSIGKSPDS 19
 || :| | :| | |
 DB 148 PENQRQPEPRDVGKVYDS 165

Search completed: August 11, 2003, 08:36:28
 Job time: 1.897 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:19:00 : Search time 26.706 seconds
(without alignments)
619.216 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103

Sequence: 1 DPEETPEVNSIGKSPDS 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 5580241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCURUS_COMB.pcp.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pcp.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pcp.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pcp.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pcp.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pcp.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pcp.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pcp.*
31: /cgn2_6/ptodata/1/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	19	26	US-10-000-039-3
2	103	100.0	82	24	US-09-937-009A-18
					Sequence 3, Appli
					Sequence 18, Appli

3	103	100.0	340	28	US-10-217-155A-17	Sequence 17, Appli
4	103	100.0	340	28	US-10-217-555-17	Sequence 17, Appli
5	103	100.0	340	28	US-10-217-555A-17	Sequence 17, Appli
6	103	100.0	340	28	US-10-217-574-17	Sequence 17, Appli
7	103	100.0	341	22	US-09-760-446A-1611	Sequence 1611, Ap
8	103	100.0	341	28	US-10-206-664-1611	Sequence 1611, Ap
9	103	100.0	388	20	US-09-646-673A-104	Sequence 104, App
10	103	100.0	388	27	US-10-131-410-104	Sequence 104, App
11	103	100.0	407	1	PCT-US03-03967-4	Sequence 4, Appli
12	103	100.0	407	26	US-10-067-977-4	Sequence 363, App
13	103	100.0	431	1	PCT-US03-01981-363	Sequence 12, Appli
14	103	100.0	431	1	PCT-US03-02571-12	Sequence 2, Appli
15	103	100.0	431	14	US-09-064-253-2	Sequence 2, Appli
16	103	100.0	431	18	US-09-405-622-2	Sequence 2, Appli
17	103	100.0	431	21	US-09-716-867-2	Sequence 2, Appli
18	103	100.0	431	22	US-09-791-537-10477	Sequence 10477, A
19	103	100.0	431	22	US-09-791-537-42519	Sequence 42519, A
20	103	100.0	431	23	US-09-810-808-5	Sequence 5, Appli
21	103	100.0	431	25	US-09-981-353-7	Sequence 2, Appli
22	103	100.0	431	26	US-10-000-039-2	Sequence 2, Appli
23	103	100.0	431	28	US-10-293-017-42	Sequence 42, Appli
24	103	100.0	431	28	US-10-293-071-42	Sequence 42, Appli
25	103	100.0	431	31	US-60-239-841-7	Sequence 7, Appli
26	103	100.0	431	31	US-60-350-061-277	Sequence 277, App
27	103	100.0	431	31	US-60-353-224-12	Sequence 12, Appli
28	103	100.0	432	22	US-09-760-443-1409	Sequence 1409, Ap
29	103	100.0	432	22	US-09-760-446A-1902	Sequence 1902, Ap
30	103	100.0	432	22	US-09-760-476-1539	Sequence 1539, Ap
31	103	100.0	432	28	US-10-206-664-1902	Sequence 1902, Ap
32	103	100.0	432	28	US-10-212-054-1409	Sequence 1409, Ap
33	103	100.0	432	28	US-10-216-245-1539	Sequence 1539, Ap
34	103	100.0	445	1	PCT-US03-03967-2	Sequence 2, Appli
35	103	100.0	445	26	US-10-067-977-2	Sequence 2, Appli
36	103	100.0	526	26	US-10-094-749-1861	Sequence 1861, Ap
37	99	96.1	434	22	US-09-791-537-29034	Sequence 29034, A
38	97	94.2	433	22	US-09-791-537-64803	Sequence 64803, A
39	97	94.2	594	22	US-09-791-537-64805	Sequence 64805, A
40	95	92.2	430	22	US-09-791-537-66219	Sequence 66219, A
41	95	92.2	430	23	US-09-810-808-9	Sequence 9, Appli
42	95	92.2	431	22	US-09-791-537-31814	Sequence 31814, A
43	95	92.2	431	22	US-09-791-537-37792	Sequence 37792, A
44	95	92.2	431	22	US-09-791-537-108642	Sequence 108642, A
45	64	62.1	52	31	US-60-162-245-2469	Sequence 2469, Ap

ALIGNMENTS

RESULT 1

US-10-000-039-3

Sequence 3, Application US/10000039

GENERAL INFORMATION:

APPLICANT: LANG, Florian

WALDEGGER, Tübingen

TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SCK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY & LANDNER

STREET: 3000 K Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/000,039

FILING DATE: 04-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/031,295
FILING DATE: 26-FEB-1998
APPLICATION NUMBER: DE 197-08-173.8
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sandercoc, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 038315/0123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-000-039-3

Query Match 100.0%; Score 103; DB 26; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
DB 1 DPEETEEPVPNSIGKSPDS 19

RESULT 2
US-09-937-009A-18
Sequence 18, Application US/09937009A
GENERAL INFORMATION:
APPLICANT: Alessi, Dario
APPLICANT: Balendran, Anudharan
APPLICANT: Deak, Maria
APPLICANT: Curtile, Richard
APPLICANT: Downes, Peter
APPLICANT: Casamayor, Antonio
TITLE OF INVENTION: Enzyme
FILE REFERENCE: 002.00170
CURRENT APPLICATION NUMBER: US/09/937,009A
CURRENT FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: PCT/GB00/01004
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: GB 9906245.7
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 82
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: region of SGK
US-09-937-009A-18

Query Match 100.0%; Score 103; DB 24; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DPEETEEPVPNSIGKSPDS 19
DB 40 DPEETEEPVPNSIGKSPDS 58

RESULT 3
US-10-217-155A-17
Sequence 17, Application US/10217155A
GENERAL INFORMATION:
APPLICANT: Barford, David
APPLICANT: Yang, Jing

APPLICANT: Hemmings, Brian A
APPLICANT: Cron, Peter D
TITLE OF INVENTION: Kinase Crystal Structures and Materials and Methods for
TITLE OF INVENTION: Kinase Activation
FILE REFERENCE: 44236
CURRENT APPLICATION NUMBER: US/10/217,155A
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: GB 0119860.5
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: GB 0209985.1
PRIOR FILING DATE: 2002-05-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 340
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Sequence source
OTHER INFORMATION: uncertain
US-10-217-155A-17

Query Match 100.0%; Score 103; DB 28; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
DB 295 DPEETEEPVPNSIGKSPDS 313

RESULT 4
US-10-217-555-17
Sequence 17, Application US/10217555
GENERAL INFORMATION:
APPLICANT: Barford, David
APPLICANT: Yang, Jing
APPLICANT: Hemmings, Brian A
APPLICANT: Cron, Peter D
TITLE OF INVENTION: Kinase Crystal Structures and Materials and Methods for
TITLE OF INVENTION: Kinase Activation
FILE REFERENCE: 44236
CURRENT APPLICATION NUMBER: US/10/217,555
CURRENT FILING DATE: 2002-12-05
PRIOR APPLICATION NUMBER: GB 0119860.5
PRIOR FILING DATE: 2001-08-14
PRIOR APPLICATION NUMBER: GB 0209985.1
PRIOR FILING DATE: 2002-05-01
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 340
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Sequence source
OTHER INFORMATION: uncertain
US-10-217-555-17

Query Match 100.0%; Score 103; DB 28; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
DB 295 DPEETEEPVPNSIGKSPDS 313

RESULT 5
US-10-217-555A-17
Sequence 17, Application US/10217555A
GENERAL INFORMATION:
APPLICANT: Barford, David


```
; APPLICANT: Yang, Jing
; APPLICANT: Hemmings, Brian A
; APPLICANT: Cron, Peter D
; TITLE OF INVENTION: Kinase Crystal Structures and Materials and Methods for
; TITLE OF INVENTION: Kinase Activation
; FILE REFERENCE: 44236
; CURRENT APPLICATION NUMBER: US/10/217,555A
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: GB 0119860.5
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: GB 0209985.1
; PRIOR FILING DATE: 2002-05-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Sequence source
; OTHER INFORMATION: uncertain
US-10-217-555A-17

Query Match 100.0%; Score 103; DB 28; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEPVPNSIGKSPDS 19
      |||||||
DB 295 DPEFTEPVPNSIGKSPDS 313

RESULT 6
US-10-217-574-17
; Sequence 17, Application US/10217574
; GENERAL INFORMATION:
; APPLICANT: Barford, David
; APPLICANT: Yang, Jing
; APPLICANT: Hemmings, Brian A
; APPLICANT: Cron, Peter D
; TITLE OF INVENTION: Kinase Crystal Structures
; FILE REFERENCE: 44237
; CURRENT APPLICATION NUMBER: US/10/217,574
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: GB 0119860.5
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: GB 0209985.1
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: GB 0216215.4
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Sequence source
; OTHER INFORMATION: uncertain
US-10-217-574-17

Query Match 100.0%; Score 103; DB 28; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEPVPNSIGKSPDS 19
      |||||||
DB 295 DPEFTEPVPNSIGKSPDS 313

RESULT 7
US-09-760-446A-1611
; Sequence 1611, Application US/09760446A
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT268
; CURRENT APPLICATION NUMBER: US/09/760,446A
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
```

; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17

; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,398

Query Match 100.0% Score 103; DB 22; Length 341;
Best Local Similarity 100.0%; Pred. No. 9,2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNISKSPDS 19
|||||
Db 296 DPEETPEVPNISKSPDS 314

RESULT 8

US-10-206-664-1611
; Sequence 1611, Application US/10206664
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ68Cin
; CURRENT APPLICATION NUMBER: US/10/206,664
; CURRENT FILING DATE: 2002-07-29
; Prior Application removed - See File Wrapper or Palm

```
; NUMBER OF SEQ ID NOS: 2628
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1611
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-206-664-1611

Query Match      100.0%; Score 103; DB 28; Length 341;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEEPVPSIGKSPDS 19
    |||||
Db 296 DPEFTEEPVPSIGKSPDS 314

RESULT 9
US-09-646-673A-104
; Sequence 104, Application US/09646673A
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/09/646,673A
; CURRENT FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-646-673A-104

Query Match      100.0%; Score 103; DB 20; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEEPVPSIGKSPDS 19
    |||||
Db 343 DPEFTEEPVPSIGKSPDS 361

RESULT 10
US-10-131-410-104
; Sequence 104, Application US/10131410
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104

; NUMBER OF SEQ ID NOS: 2628
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1611
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-104

Query Match      100.0%; Score 103; DB 27; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEEPVPSIGKSPDS 19
    |||||
Db 343 DPEFTEEPVPSIGKSPDS 361

RESULT 11
PCT-US03-03967-4
; Sequence 4, Application PC/TUS0303967
; GENERAL INFORMATION:
; APPLICANT: APPLERA CORPORATION et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001313PCT
; CURRENT APPLICATION NUMBER: PCT/US03/03967
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-03967-4

Query Match      100.0%; Score 103; DB 1; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEEPVPSIGKSPDS 19
    |||||
Db 362 DPEFTEEPVPSIGKSPDS 380

RESULT 12
US-10-067-977-4
; Sequence 4, Application US/10067977
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua and KE, Zhaoxi
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001313
; CURRENT APPLICATION NUMBER: US/10/067,977
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-977-4

Query Match      100.0%; Score 103; DB 26; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEFTEEPVPSIGKSPDS 19
    |||||
Db 362 DPEFTEEPVPSIGKSPDS 380

RESULT 13
PCT-US03-01981-363
; Sequence 363, Application PC/TUS0301981
```

```
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/01981
; PRIOR FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: Patent version 3.2
; SEQ ID NO 363
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-01981-363

Query Match 100.0%; Score 103; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEETPEVPNSIGKSPDS 19
Db 386 DPEETPEVPNSIGKSPDS 404

RESULT 14
PCT-US03-02571-12
; Sequence 12, Application PC/TUS0302571
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodriquez-Way, Amelle
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1882, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 29002, 33216, 43726, 69292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 8870, 25856, 32394, 3484, 345, 9252, 9135,
; TITLE OF INVENTION: 10332, 18610, 8165, 2448, 2445, 64624, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; TITLE OF INVENTION: 9792, 15400, 1452 or 6585 molecules
; FILE REFERENCE: MPI02-018
; CURRENT APPLICATION NUMBER: PCT/US03/02571
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
```

```
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homosapien
PCT-US03-02571-12

Query Match 100.0%; Score 103; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEETPEVPNSIGKSPDS 19
Db 386 DPEETPEVPNSIGKSPDS 404

RESULT 15
US-09-064-253-2
; Sequence 2, Application US/09064253
; GENERAL INFORMATION:
; APPLICANT: Kumar, Janet
; TITLE OF INVENTION: Human Serum Glucocorticoid
; TITLE OF INVENTION: Regulated Kinase, A Target For Chronic Renal Disease
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/064,253
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051,125
; FILING DATE: 22-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Had, William T.
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GPO0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)270-5219
; TELEFAX: (610)270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-064-253-2

Query Match 100.0%; Score 103; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEETPEVPNSIGKSPDS 19
Db 386 DPEETPEVPNSIGKSPDS 404
```

Search completed: August 11, 2003, 08:35:53
Job time : 26.706 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:22:25 ; Search time 15.412 Seconds
(without alignments)
156.930 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103
Sequence: 1 DPEFTEPVPNSICKSPDS 19

Scoring Table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 479057 seqs, 127295195 residues

Total number of hits satisfying chosen parameters: 479057

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	103	100.0	19	15	US-10-000-039-3
2	103	100.0	431	9	US-09-810-808-5
3	103	100.0	431	9	US-09-810-808-5
4	103	100.0	431	15	US-10-000-039-2
5	95	92.2	430	9	US-09-810-808-9
6	64	62.1	367	10	US-09-971-118-2
7	46	44.7	93	10	US-09-925-300-1638
8	46	44.7	142	11	US-09-764-891-3889
9	44	42.7	194	9	US-09-784-249-4
10	44	42.7	481	16	US-10-080-170-14
11	44	42.7	496	9	US-09-784-249-2
12	43	41.7	111	15	US-10-006-869-21
13	43	41.7	146	15	US-10-156-761-11658
14	43	41.7	600	11	US-09-893-519A-71
15	43	41.7	941	14	US-10-124-557-14

Sequence 84, Appl
Sequence 74, Appl
Sequence 58, Appl
Sequence 104, Appl
Sequence 44, Appl
Sequence 42, Appl
Sequence 142, Appl
Sequence 50, Appl
Sequence 46, Appl
Sequence 60, Appl
Sequence 48, Appl
Sequence 40, Appl
Sequence 52, Appl
Sequence 2, Appl
Sequence 62, Appl
Sequence 59, Appl
Sequence 59, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 20, Appl
Sequence 13068, A
Sequence 12514, A
Sequence 8, Appl
Sequence 16, Appl
Sequence 11924, A
Sequence 11834, A
Sequence 5, Appl
Sequence 21, Appl
Sequence 213, Appl

US-10-124-557-84
US-10-124-557-74
US-10-124-557-58
US-10-124-557-104
US-10-124-557-44
US-10-124-557-42
US-10-124-557-142
US-10-124-557-50
US-10-124-557-46
US-10-124-557-60
US-10-124-557-48
US-10-124-557-40
US-10-124-557-52
US-10-124-557-2
US-10-124-557-62
US-10-106-698-5932
US-10-028-245-2
US-10-024-368-6
US-10-246-354-2
US-10-246-354-3
US-09-801-574-20
US-10-156-761-13068
US-10-156-761-12514
US-10-126-279-8
US-10-160-865-16
US-10-156-761-11924
US-10-156-761-11834
US-10-024-368-5
US-10-108-603-211
US-10-108-603-213

ALIGNMENTS

RESULT 1
US-10-000-039-3

; Sequence 3, Application US/10000039
; Publication No. US2003003559A1

; GENERAL INFORMATION:

APPLICANT: LANG, Florian

TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: FOLEY & LARDNER

STREET: 3000 K Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/000,039

FILING DATE: 04-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/031,295

FILING DATE: 26-FEB-1998

APPLICATION NUMBER: DE 197-08-173.8

FILING DATE: 28-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sandercock, Colin G.

REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 058315/0123

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-000-039-3

Query Match 100.0%; Score 103; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEPVPSIGKSPDS 19
Db 1 DPEETEPVPSIGKSPDS 19
|||||

RESULT 2
US-09-810-808-5
Sequence 5, Application US/09810808
Patent No. US20020421141
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
Guegler, Karl J.
Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/810,808

FILING DATE: 15-Mar-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/541,228

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 431 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: <unknown>

CLONE: Consensus

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-810-808-5

Query Match 100.0%; Score 103; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEPVPSIGKSPDS 19
Db 386 DPEETEPVPSIGKSPDS 404
|||||

RESULT 3
US-09-981-353-7
Sequence 7, Application US/09981353
Patent No. US20020160382A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy W.
APPLICANT: Jones, David A.
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
FILE REFERENCE: PA-0038 US
CURRENT APPLICATION NUMBER: US/09/981,353
CURRENT FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 7
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020160382A1 3819039CD1
US-09-981-353-7

Query Match 100.0%; Score 103; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEPVPSIGKSPDS 19
Db 386 DPEETEPVPSIGKSPDS 404
|||||

RESULT 4
US-10-000-039-2
Sequence 2, Application US/10000039
Publication No. US2003003559A1
GENERAL INFORMATION:
APPLICANT: LANG, Florian
TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SCK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/000,039
FILING DATE: 04-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/031,295
FILING DATE: 26-FEB-1998
APPLICATION NUMBER: DE 197-08-173.8
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Collin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 058315/0123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-000-039-2

Query Match 100.0%; Score 103; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSICKSPDS 19
IIIIIIIIIIIIIIIIIIII
DB 386 DPEETEEVPVNSICKSPDS 404

RESULT 5

US-09-810-808-9
; Sequence 9, Application US/09810808
; Patent No. US20020042114A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; Guegler, Karl J.
; Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/810.808
FILING DATE: 15-Mar-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/541,228
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-810-808-9

Query Match 92.2%; Score 95; DB 9; Length 430;
Best Local Similarity 89.5%; Pred. No. 3.5e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSICKSPDS 19
IIIIIIIIIIIIIIIIIIII
DB 385 DPEETEEVPVNSICKSPDS 403

RESULT 6

US-09-971-118-2
; Sequence 2, Application US/09871118
; Patent No. US20020123056A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILAINATHAN
; TITLE OF INVENTION: SGK2 AND ITS USES
; FILE REFERENCE: KINE025CIP
; CURRENT APPLICATION NUMBER: US/09/971.118
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/21479
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/237,419
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-971-118-2

Query Match 62.1%; Score 64; DB 10; Length 367;
Best Local Similarity 63.2%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSICKSPDS 19
IIIIIIIIIIIIIIIIIIII
DB 323 DPEETEEVPVNSICKSPDS 341

RESULT 7

US-09-925-300-1638
; Sequence 1638, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925.300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1638
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (90)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1638

Query Match 44.7%; Score 46; DB 10; Length 93;
Best Local Similarity 47.1%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSICKSP 17
IIIIIIIIIIIIIIIIIIII
DB 42 DPEETEEVPVNSICKSP 58

RESULT 8

US-09-764-891-3889
; Sequence 3889, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

```

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3889
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (118)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (122)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3889

```

```

Query Match 44.7%; Score 46; DB 11; Length 142;
Best Local Similarity 44.4%; Pred. No. 19;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 2 PEETEEVPVNSIGKSPD 19
Db 80 PSTTHRPVPLVGVGPEDT 97

```

```

RESULT 9
US-09-784-249-4
; Sequence 4, Application US/09784249
; Patent No. US20010027184A1
; GENERAL INFORMATION:
; APPLICANT: Sanjay Kumar
; APPLICANT: Cheng Zou
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE (H-SGRK2)
; FILE REFERENCE: GH-70124-C1
; CURRENT APPLICATION NUMBER: US/09/784,249
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 08/997,212
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/051,446
; PRIOR FILING DATE: 1997-07-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 194
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-784-249-4

```

```

Query Match 42.7%; Score 44; DB 9; Length 194;
Best Local Similarity 55.6%; Pred. No. 53;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 DPEETEEVPVNSIGKSPD 18
Db 148 DTAFTETVPVSVCVSSD 165

```

```

RESULT 10
US-10-080-170-14
; Sequence 14, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TREATMENT OF MYCOBACTERIOSES

```

```

; FILE REFERENCE: 03495, 0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-080-170-14

```

```

Query Match 42.7%; Score 44; DB 16; Length 481;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 2 PEETEEVPVNSIGKSPD 18
Db 245 PTIPDTGAPNSLGLSPD 261

```

```

RESULT 11
US-09-784-249-2
; Sequence 2, Application US/09784249
; Patent No. US20010027184A1
; GENERAL INFORMATION:
; APPLICANT: Sanjay Kumar
; APPLICANT: Cheng Zou
; TITLE OF INVENTION: SERINE/THREONINE PROTEIN KINASE (H-SGRK2)
; FILE REFERENCE: GH-70124-C1
; CURRENT APPLICATION NUMBER: US/09/784,249
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 08/997,212
; PRIOR FILING DATE: 1997-12-23
; PRIOR APPLICATION NUMBER: 60/051,446
; PRIOR FILING DATE: 1997-07-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-784-249-2

```

```

Query Match 42.7%; Score 44; DB 9; Length 496;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 DPEETEEVPVNSIGKSPD 18
Db 450 DTAFTETVPVSVCVSSD 467

```

```

RESULT 12
US-10-006-869-21
; Sequence 21, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens

```


US-10-006-869-21

Query Match 41.7% Score 43; DB 15; Length 111;
Best Local Similarity 44.4%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PFTTEPVPNSIGKSPDS 19
II : I : II II
Db 14 PENORQPPRDVGKVVDS 31

RESULT 13

US-10-156-761-11658
; Sequence 11658, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11658
; LENGTH: 146
; TYPE: PRT

; ORGANISM: Streptomyces avermitilis
US-10-136-761-11658

Query Match 41.7% Score 43; DB 15; Length 146;
Best Local Similarity 57.1%; Pred. No. 55;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 TEEPVPNSIGKSPD 18
IIII I:: III
Db 118 TEPTPSAEPSPD 131

RESULT 14

US-09-893-519A-71
; Sequence 71, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey

; APPLICANT: BUORMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOWARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene

; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/1G548-US2
; CURRENT APPLICATION NUMBER: US/09/893.519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164

; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 144
US-09-893-519A-71

Query Match 41.7% Score 43; DB 11; Length 600;
Best Local Similarity 53.3%; Pred. No. 2.6e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 TEEPVPNSIGKSPDS 19
:IIII I II: I
Db 43 SDEVPESAGKAUTS 57

RESULT 15

US-10-124-557-14
; Sequence 14, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.

; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989

; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 941 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-124-557-14

Query Match 41.7%; Score 43; DB 14; Length 941;
Best Local Similarity 43.8%; Pred. No: 4.2e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PFTTEPVPNSIGKSP 17
|| : || | : ||
Db 679 PELSAEPTPKALENSP 694

Search completed: August 11, 2003, 08:42:59
Job time : 16.412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 07:58:29 ; Search time 5.09657 Seconds
(without alignments)
962.020 Million cell updates/sec

Title: US-10-000-039a-3

Perfect score: 103

Sequence: 1 DPEFTEPVPNSIGKSPDS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp-archaea.*

2: sp-bacteria.*

3: sp-fungi.*

4: sp-human.*

5: sp-invertebrate.*

6: sp-mammal.*

7: sp-mhc.*

8: sp-organelle.*

9: sp-phage.*

10: sp-plant.*

11: sp-rodent.*

12: sp-virus.*

13: sp-vertebrate.*

14: sp-unclassified.*

15: sp-rvivirus.*

16: sp-bacteriap.*

17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	96.1	434	13 O93524	O93524 xenopus lae
2	97	94.2	433	13 O73926	O73926 squalus aca
3	97	94.2	594	13 O73927	O73927 squalus aca
4	95	92.2	185	11 O99104	O99104 mus musculus
5	96	94.1	40	6 O95N98	O95N98 canis famli
6	48	46.6	448	10 Q9LX59	Q9LX59 arabidopsis
7	48	46.6	490	16 Q8UA45	Q8UA45 agrobacteri
8	47.5	46.1	163	10 Q9SL25	Q9SL25 arabidopsis
9	47.5	46.1	4006	11 O35452	O35452 mus musculus
10	47.5	46.1	4114	11 O34796	O34796 mus musculus
11	47	45.6	178	5 Q8IRG0	Q8IRG0 drosophila
12	47	45.6	311	2 Q8GAF7	Q8GAF7 arthrobacte
13	47	45.6	490	16 Q84052	Q84052 chlamydia t
14	47	45.6	494	17 Q96YCB	Q96YCB sulfolobu
15	47	45.6	768	5 Q8SRC6	Q8SRC6 encephalico
16	46	44.7	315	4 Q8NWB5	Q8NWB5 homo sapien

17 46 44.7 747 16 Q8U189
18 46 44.7 1965 16 Q8XLJ5
19 45.5 44.2 1301 4 Q9COC1
20 45 43.7 145 12 Q8JKY3
21 45 43.7 149 11 Q9CU16
22 45 43.7 243 2 Q31361
23 45 43.7 336 10 Q9FT05
24 45 43.7 339 10 Q9ZQM9
25 45 43.7 446 13 Q8AXX8
26 45 43.7 464 13 Q8JH47
27 45 43.7 467 17 Q8T230
28 45 43.7 548 5 Q44650
29 45 43.7 708 10 Q9LGM8
30 45 43.7 843 11 Q3JJB4
31 45 43.7 1555 5 Q8T1Q4
32 45 43.7 1654 4 Q9PIY6
33 44.5 43.2 524 10 Q9SBN2
34 44.5 43.2 616 10 Q9FY49
35 44.5 43.2 858 11 Q55157
36 44.5 43.2 1339 11 Q35788
37 44 42.7 144 10 Q64575
38 44 42.7 230 13 Q8AVP7
39 44 42.7 260 11 Q9D521
40 44 42.7 260 11 Q8BPE1
41 44 42.7 310 11 Q9CSM2
42 44 42.7 315 11 Q8CHP9
43 44 42.7 330 5 Q967S9
44 44 42.7 481 16 Q33088
45 44 42.7 640 3 Q02495

ALIGNMENTS

RESULT 1

O93524 ID O93524 PRELIMINARY; PRT; 434 AA.

AC O93524;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 49.1 kDa protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Renal tubule;
RX MEDLINE=99162637; PubMed=10051674;
RA Chen S.Y., Bhargava A., Mastroberardino L., Meijer O.C., Wang J.,
RA Buse P., Firestone G.L., Verrey F., Pearce D.;
RT "Epithelial sodium channel regulated by aldosterone-induced protein
RT sgk.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2514-2519(1999).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF057138; ARC62398.1;
DR HSP: P00517; IYDR.
DR InterPro: IPR000961; pkinase_C.
DR InterPro: IPR000719; prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase_C; 1.
DR ProDom: PD0000001; Prot_Kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00133; S_TK X; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 434 AA: 49130 MW: 4A061E38B6AA6F61 CRC64;

O8U189 agrobacteri
Q8XLJ5 clostridium
Q9COC1 homo sapien
Q8JKY3 virus phich
Q9CU16 m adult mal
Q31361 borrelia ga
Q9FT05 cicer ariet
Q9ZQM9 arabidopsis
Q8AXX8 xenopus lae
Q8JH47 brachydanio
Q8T230 methanopyru
Q44650 caenorhabdi
Q9LGM8 oryza sativ
Q3JJB4 mus musculu
Q8T1Q4 trypanosoma
Q9PIY6 homo sapien
Q9SBN2 volvox carl
Q9FY49 arabidopsis
Q55157 rattus norv
Q35788 rattus norv
Q64575 arabidopsis
Q8AVP7 xenopus lae
Q9D521 mus musculu
Q8BPE1 mus musculu
Q9CSM2 mus musculu
Q8CHP9 mus musculu
Q967S9 toxoplasma
Q33088 mycobacteri
Q02495 schizophy11

```
Query Match          96.1%; Score 99; DB 13; Length 434;
Best Local Similarity 94.7%; Pred. No. 1.3e-07;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
    |||||.....|
Db 389 DPEETEEPVPNSIGKSPDS 407

RESULT 2
O73926 PRELIMINARY; PRT; 433 AA.
AC O73926;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE S-sqk1.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidel; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN 11
RP SEQUENCE FROM N.A.
RA Waldegger S., Barth P., Forrest J., Greger R.F., Lang F.;
RT "Cloning of sgk Serine-threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AJ223715; CAA11527.1;
DR HSP: P05132; IFWO.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00133; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 433 AA; 48369 MW; 982BD01A59AA3A78 CRC64;

Query Match          94.2%; Score 97; DB 13; Length 433;
Best Local Similarity 89.5%; Pred. No. 2.7e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
    |||||.....|
Db 388 DPEETEEPVPNSIGKSPDS 406

RESULT 3
O73927 PRELIMINARY; PRT; 594 AA.
AC O73927;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE S-sqk2.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidel; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN 11
RP SEQUENCE FROM N.A.
RA Waldegger S., Barth P., Forrest J., Greger R.F., Lang F.;
RT "Cloning of sgk Serine-threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
EMBL: AJ223716; CAA11528.1;
HSP: P00517; 1YDR.
InterPro: IPR000961; Pkinase_C.
InterPro: IPR002290; Ser_thr_pkinase.
Pfam: PF00069; pkinase; 1.
ProDom: PD000001; Prot_kinase; 1.
SMART: SM00220; S_TKc; 1.
SMART: SM00133; S_TKc; 1.
PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 594 AA; 66857 MW; 40153ECL8D57B767 CRC64;

Query Match          94.2%; Score 97; DB 13; Length 594;
Best Local Similarity 89.5%; Pred. No. 3.9e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
    |||||.....|
Db 549 DPEETEEPVPNSIGKSPDS 567

RESULT 4
O99LU04 PRELIMINARY; PRT; 185 AA.
AC O99LU04;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to serum/glucocorticoid regulated kinase (Fragment).
GN SGK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC002222; AA02222.1;
DR MGD; MGI:1340062; Sgk.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase_C; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00133; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 185 AA; 20921 MW; EDAAEA44ABF083945 CRC64;

Query Match          92.2%; Score 95; DB 11; Length 185;
Best Local Similarity 89.5%; Pred. No. 2.3e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGKSPDS 19
    |||||.....|
Db 140 DPEETEEPVPNSIGKSPDS 158

RESULT 5
O95N98 PRELIMINARY; PRT; 40 AA.
AC O95N98;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Serum/glucocorticoid-regulated kinase 1 (Fragment).
GN SGK1.
```

OS Canis familiaris (Dog);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21168084; PubMed=11266509;
RA Nick V.E., Itani O.A., Loftus R.W., Husted R.F., Schmidt T.J.,
Thomas C.F.;
RA "The alpha-Subunit of the Epithelial Sodium Channel Is an Aldosterone-
Induced Transcript in Mammalian Collecting Ducts, and This
Transcriptional Response Is Mediated via Distinct cis-Elements in the
5'-Flanking Region of the Gene";
RL Mol. Endocrinol. 15:575-588(2001).
DR EMBL: AF317416; AAK54044.1;
KW Kinase.
FT NON_TER
SQ SEQUENCE 40 AA; 4189 MW; E54EB18676E225B6 CRC64;

Query Match 64.1%; Score 66; DB 6; Length 40;
Best Local Similarity 92.3%; Pred. No. 0.0021;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 EPVPNSIGKSPDS 19
Db 1 EPVPNSIGKSPDS 13

RESULT 6
Q9LX59
ID Q9LX59 PRELIMINARY; PRT; 448 AA.
AC Q9LX59;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE Hypothetical 50.9 kDa protein.
GN F4M19_70.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Vitale D., Liquori R., Flores M., Argiriou A., De Simone V.,
Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
Salanoubat M.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL356013; CAB91582.1;
DR InterPro: IPR003871; DUF223.
DR Pfam: PF02721; DUF223; 1.
KW Hypothetical protein.
SQ SEQUENCE 448 AA; 50924 MW; 5421BFBB477B1FB CRC64;

Query Match 46.6%; Score 48; DB 10; Length 448;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DPEFTEPVPNSIGKSPDS 16
Db 315 DPELPEPVKNLIGKT 330

RESULT 7
Q8UA45
ID Q8UA45 PRELIMINARY; PRT; 490 AA.
AC Q8UA45;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Mannitol dehydrogenase.
GN XUB OR ATU5330 OR AGR_L2604.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Woo L.,
Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
Kutyavin T., Levy R., Li M.-J., McLellan E., Palmeri A.,
Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Bolan M.,
Chumley F., Tingey S.V., Tomb J.F., Gordon M.P., Olson M.V.,
Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
C58";
RT Science 294:2317-2323(2001).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
Ourolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
Roumiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
Agrobacterium tumefaciens C58";
RT Science 294:2323-2328(2001).
RL EMBL: AE009282; AAL44342.1;
DR EMBL: AE008330; AAK89867.1;
DR InterPro: IPR000669; Mannitol_dh.
DR Pfam: PF01232; Mannitol_dh; 1.
DR PRINTS: PR00084; MTLDHGRNASE.
KW Complete proteome.
SQ SEQUENCE 490 AA; 53904 MW; 374B62F989CADD07 CRC64;

Query Match 46.6%; Score 48; DB 16; Length 490;
Best Local Similarity 47.4%; Pred. No. 26;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DPEFTEPVPNSIGKSPDS 19
Db 423 DPAFAEMAVPTGNGEPPDA 441

RESULT 8
Q9SL25
ID Q9SL25 PRELIMINARY; PRT; 163 AA.
AC Q9SL25;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE At2g09840 protein.
GN AT2G09840.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RL MEDLINE-20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fujii C.Y., Mason T.M., Bowman C.H., Barnstead M.E., Feldblum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

[illegible]

characterization of 2,6-dihydroxypyridine 3-hydroxylase.";
J. Bacteriol. 183:5262-5267(2001).
[7]
SEQUENCE FROM N.A.
RA Igloi G.B., Brandisch R.; Catabolic Plasmid pAO1 from *Arthrobacter*
RT "Sequence of the 105 kb catabolic plasmid of a pAO1-dependent Nicotine Uptake
RT nicotinovorans and identification of a pAO1-dependent Nicotine Uptake
RT system".
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ507836; CAD47974.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 311 AA; 3928 MW; 0C00102DC0DB6D53 CRC64;

Query Match 45.6%; Score 47; DB 2; Length 311;
Best Local Similarity 53.3%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 PFTPEPVPNSIGSKS 16
| : : : : :
Db 10 PSPTQPPVPDIVGKS 24

RESULT 13
ID 084052 PRELIMINARY; PRT; 490 AA.
AC 084052;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DF 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CT049.
GN CT049.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UV-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Ollinger L., Ratsov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis";
RL Science 282:754-759(1998).
DR EMBL: AE001279; AAC67640.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 490 AA; 50733 MW; 533280EC9D4DFEE CRC64;

Query Match 45.6%; Score 47; DB 16; Length 490;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 PFTPEPVPNSIGSKSPDS 19
| : : : : :
Db 318 PFWTQGEVPLVEQSPDS 335

RESULT 14
ID 096YC8 PRELIMINARY; PRT; 494 AA.
AC 096YC8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DF 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein ST2240.
GN ST2240.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;

PubMed=11572479;
RA Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anaki A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT crenarchaeon, *Sulfolobus tokodaii* strain 7".
RL DNA Res. 8:123-140(2001).
DR EMBL: AF000989; BAB67349.1; -
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 494 AA; 57555 MW; 863A2013340A6E0D CRC64;

Query Match 45.6%; Score 47; DB 17; Length 494;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 FTEEPVPNSIGSKS 16
| : : : : :
Db 86 EFSEEEVQKSIGKS 99

RESULT 15
ID 08SRC6 PRELIMINARY; PRT; 768 AA.
AC 08SRC6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DF 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BELONGS to the XPF/ERCC4/RAD1 family.
GN ECU08_0760.
OS Eucaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OC Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX NCBI_TaxID=6035;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RA Genoscope;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat P.,
RA Prensier G., Barbe V., Peyretailade E., Brottier P., Wincker P.,
RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Encephalitozoon cuniculi".
RL Nature 414:450-453(2001).
DR EMBL: AL590448; CAD26381.1; -
DR InterPro: IPR006166; ERCC4.
DR Pfam: PF02732; ERCC4; 1.
SQ SEQUENCE 768 AA; 87858 MW; B05F46C01679DA15 CRC64;

Query Match 45.6%; Score 47; DB 5; Length 768;
Best Local Similarity 64.3%; Pred. No. 62;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 FTEEPVPNSIGSKSP 17
| : : : : :
Db 180 FSESVPVLSIGFSP 193

Search completed: August 11, 2003, 08:22:14
Job time : 9.09657 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:14:59 ; Search time 2.69099 Seconds
(without alignments)
298.740 Million cell updates/sec

Title: US-10-000-039A-3

Perfect score: 103
Sequence: 1 DPETPEPVNSICKSPDS 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2.6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2.6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2.6/prodata/1/1aa/6A-COMB.pep:*
4: /cgn2.6/prodata/1/1aa/6B-COMB.pep:*
5: /cgn2.6/prodata/1/1aa/FCUS-COMB.pep:*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	103	100.0	19	US-09-031-295-3	Sequence 3, Appli
2	103	100.0	431	US-08-712-709-5	Sequence 5, Appli
3	103	100.0	431	US-09-111-444-5	Sequence 5, Appli
4	103	100.0	431	US-09-541-228-5	Sequence 5, Appli
5	103	100.0	431	US-09-031-295-2	Sequence 2, Appli
6	95	92.2	430	US-08-712-709-9	Sequence 9, Appli
7	95	92.2	430	US-09-111-444-9	Sequence 9, Appli
8	95	92.2	430	US-09-541-228-9	Sequence 9, Appli
9	44	42.7	540	US-09-134-001C-4705	Sequence 21, Appl
10	43	41.7	111	US-09-187-859-21	Sequence 21, Appl
11	43	41.7	111	US-09-839-542B-21	Sequence 21, Appl
12	43	41.7	559	US-08-313-288B-14	Sequence 14, Appl
13	43	41.7	687	US-08-540-804-4	Sequence 4, Appli
14	43	41.7	687	US-08-218-265-4	Sequence 4, Appli
15	43	41.7	687	US-08-521-872-4	Sequence 4, Appli
16	43	41.7	687	US-08-590-399-4	Sequence 4, Appli
17	43	41.7	713	US-08-188-228-62	Sequence 62, Appl
18	43	41.7	713	US-08-332-643-56	Sequence 56, Appl
19	43	41.7	713	US-08-332-638-62	Sequence 62, Appl
20	43	41.7	941	US-07-757-022B-14	Sequence 14, Appl
21	43	41.7	1022	US-07-757-022B-84	Sequence 84, Appl
22	43	41.7	1038	US-07-757-022B-74	Sequence 74, Appl
23	43	41.7	1049	US-07-757-022B-58	Sequence 58, Appl
24	43	41.7	1140	US-07-757-022B-104	Sequence 104, App
25	43	41.7	1270	US-07-757-022B-44	Sequence 44, Appl
26	43	41.7	1311	US-07-757-022B-42	Sequence 42, Appl
27	43	41.7	1313	US-07-757-022B-142	Sequence 142, App

28	43	41.7	1314	4	US-07-757-022B-50	Sequence 50, Appl
29	43	41.7	1320	4	US-07-757-022B-46	Sequence 46, Appl
30	43	41.7	1320	4	US-07-757-022B-60	Sequence 60, Appl
31	43	41.7	1354	4	US-07-757-022B-48	Sequence 48, Appl
32	43	41.7	1361	4	US-07-757-022B-40	Sequence 40, Appl
33	43	41.7	1363	4	US-07-757-022B-52	Sequence 52, Appl
34	43	41.7	1404	4	US-07-757-022B-2	Sequence 2, Appli
35	43	41.7	1404	4	US-07-757-022B-62	Sequence 62, Appl
36	42	40.8	57	2	US-08-619-542B-39	Sequence 39, Appl
37	42	40.8	176	2	US-08-619-542B-44	Sequence 44, Appl
38	42	40.8	464	2	US-08-969-630-5	Sequence 5, Appli
39	42	40.8	548	4	US-09-107-532A-6627	Sequence 6627, Ap
40	42	40.8	1182	4	US-09-287-354-6	Sequence 6, Appli
41	42	40.8	3052	2	US-08-557-122A-26	Sequence 26, Appl
42	42	40.8	3052	4	US-09-262-666-26	Sequence 26, Appl
43	41	39.8	328	4	US-09-252-991A-30671	Sequence 30671, A
44	41	39.8	601	2	US-08-795-868-16	Sequence 16, Appl
45	41	39.8	602	4	US-09-303-069-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-031-295-3
; Sequence 3, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGER, Tubingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SCK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: DE 197-08-173.8
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-031-295-3

Query Match 100.0%; Score 103; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DPETPEPVNSICKSPDS 19
|||||

Db 1 DPEETEPVPSIGKSPDS 19

RESULT 2

US-08-712-709-5
; Sequence 5, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE: Filed Herewith
; APPLICATION NUMBER: US/08/712,709
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus

US-08-712-709-5
Query Match 100.0%; Score 103; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 7,1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEETEPVPSIGKSPDS 19

Db 386 DPEETEPVPSIGKSPDS 404

RESULT 3

US-09-111-444-5
; Sequence 5, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/541,228
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/111,444
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus

US-09-111-444-5
Query Match 100.0%; Score 103; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 7,1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEETEPVPSIGKSPDS 19

Db 386 DPEETEPVPSIGKSPDS 404

RESULT 4

US-09-541-228-5
; Sequence 5, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/541,228
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```
;; TELEFAX: 415-845-4166
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 431 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; IMMEDIATE SOURCE:
;; LIBRARY:
;; CLONE: Consensus
US-09-541-228-5
Query Match 100.0%; Score 103; DB 3; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSIGKSPDS 19
|||||
DB 386 DPEETEEVPVNSIGKSPDS 404

RESULT 5
US-09-031-295-2
; Sequence 2, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGGER, Tubingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-295-2
Query Match 100.0%; Score 103; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSIGKSPDS 19
|||||
DB 386 DPEETEEVPVNSIGKSPDS 404
```

```
;; TELEFAX: 415-845-4166
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 431 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; IMMEDIATE SOURCE:
;; LIBRARY:
;; CLONE: Consensus
US-08-712-709-9
Query Match 92.2%; Score 95; DB 2; Length 430;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSIGKSPDS 19
|||||
DB 385 DPEETEEVPVNSIGKSPDS 403

RESULT 6
US-08-712-709-9
; Sequence 9, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PR-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 294637
US-08-712-709-9
Query Match 92.2%; Score 95; DB 2; Length 430;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEVPVNSIGKSPDS 19
|||||
DB 385 DPEETEEVPVNSIGKSPDS 403

RESULT 7
US-09-111-444-9
; Sequence 9, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA: US/09/111,444
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637
US-09-111-444-9

Query Match 92.2%; Score 95; DB 3; Length 430;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGRSPDS 19
|||||:|||||:|||||
DB 385 DPEETEEPVPNSIGRSPDS 403

RESULT 8
US-09-541-228-9
Sequence 9, Application US/09541228
Patent No. 6232077
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jncvte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637
US-09-541-228-9

Query Match 92.2%; Score 95; DB 3; Length 430;
Best Local Similarity 89.5%; Pred. No. 1.2e-06;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGRSPDS 19
|||||:|||||:|||||
DB 385 DPEETEEPVPNSIGRSPDS 403

RESULT 9
US-09-134-001C-4705
Sequence 4705, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4705
LENGTH: 540
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4705

Query Match 42.7%; Score 44; DB 4; Length 540;
Best Local Similarity 47.1%; Pred. No. 76;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 DPEETEEPVPNSIGRSP 17
|||:|:|:|:|
DB 356 DSEPKKXITESLGKLP 372

RESULT 10
US-09-187-859-21
Sequence 21, Application US/09187859A
Patent No. 6358920
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
FILE REFERENCE: 100086.407C1
CURRENT APPLICATION NUMBER: US/09/187,859A
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 4052
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 111
TYPE: PRT
ORGANISM: Homo sapiens
US-09-187-859-21

Query Match 41.7%; Score 43; DB 4; Length 111;
Best Local Similarity 44.4%; Pred. No. 19;

Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 2 PEFTEPVPNSIGKSPDS 19
11 : 1 : 1 : 1 : 1 : 1 :
Db 14 PENQRPFRDVGKVD 31

RESULT 11

US-09-839-542B-21
; Sequence 21, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Cour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-542B-21

Query Match 41.7%; Score 43; DB 4; Length 111;
Best Local Similarity 44.4%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 2 PEFTEPVPNSIGKSPDS 19
11 : 1 : 1 : 1 : 1 : 1 :
Db 14 PENQRPFRDVGKVD 31

RESULT 12

US-08-313-288B-14
; Sequence 14, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 559 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-313-288B-14

Query Match 41.7%; Score 43; DB 1; Length 559;
Best Local Similarity 41.2%; Pred. No. 1.1e+02;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PEFTEPVPNSIGKSPD 18
11 : 1 : 1 : 1 : 1 : 1 :
Db 365 PEDSEKEVPSPDVKPNE 181

RESULT 13

US-08-540-804-4
; Sequence 4, Application US/08540804
; Patent No. 5919666
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Chao, David M.
; TITLE OF INVENTION: No. 5919666el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Therefor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,804
; FILING DATE: 11-OCT-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/521,872
; FILING DATE: 21-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH194-03A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-540-804-4

Query Match 41.7%; Score 43; DB 2; Length 687;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TEEPVPNSIGKSPDS 19
11 : 1 : 1 : 1 : 1 : 1 :
Db 43 SDEVPESAGKADTS 57

RESULT 14
US-08-218-265-4
; Sequence 4, Application US/08218265
; Patent No. 5922585
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; TITLE OF INVENTION: No. 5922585el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,265
; FILING DATE: 25-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-218-265-4

Query Match 41.7%; Score 43; DB 2; Length 687;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 5 TEEPVPNSIGKSPDS 19
Db 43 SDEVPESAGKADTS 57
::|||||::|

RESULT 15
US-08-521-872-4
; Sequence 4, Application US/08521872
; Patent No. 6015682
; GENERAL INFORMATION:
; APPLICANT: Young, Richard A.
; APPLICANT: Koleske, Anthony J.
; APPLICANT: Thompson, Craig M.
; APPLICANT: Chao, David M.
; TITLE OF INVENTION: No. 6015682el Factors Which Modify Gene
; TITLE OF INVENTION: Transcription and Methods of Use Therefor
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/521,872
; FILING DATE: 31-AUG-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/218,265
; FILING DATE: 25-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-03A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 687 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-521-872-4
Query Match 41.7%; Score 43; DB 3; Length 687;
Best Local Similarity 53.3%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 5 TEEPVPNSIGKSPDS 19
Db 43 SDEVPESAGKADTS 57
::|||||::|
Search completed: August 11, 2003, 08:24:44
Job time : 3.69099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:11:03 ; Search time 1.87554 Seconds
(without alignments)
476.401 Million cell updates/sec

Title: US-10-000-039A-3

Sequence: 1 DPETEPVPNSIGKSPDS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	431	1 SGK1_HUMAN	O00141 homo sapien
2	95	92.2	430	1 SGK1_RAT	O05226 rattus norv
3	95	92.2	431	1 SGK1_MOUSE	O99vc6 mus musculu
4	95	92.2	431	1 SGK1_RABIT	O9xt18 oryctolagus
5	64	62.1	367	1 SGK2_MOUSE	O9qz55 mus musculu
6	64	62.1	427	1 SGK2_HUMAN	O9hby8 homo sapien
7	58	56.3	460	1 VL2_HPV55	O80939 human papil
8	51	49.5	620	1 EXTN_TOBAC	P13983 nicotiana t
9	47	45.6	287	1 UPK3_HUMAN	O75631 homo sapien
10	46.5	45.1	1664	1 SLPI_CLOTH	O08852 clostridium
11	46	44.7	233	1 TNFA_CANFA	P51742 canis fami
12	46	44.7	287	1 UPK3_BOVIN	P38574 bos taurus
13	46	44.7	464	1 SPN5_SCHPO	P48010 schizosacch
14	46	44.7	837	1 YLF5_CAEEL	P46941 caenorhabdi
15	45	43.7	325	1 BMPA_BORGA	O31357 borrelia ga
16	45	43.7	374	1 YDB1_SCHPO	O10354 schizosacch
17	44	42.7	249	1 CISH_CHICK	O99w70 gallus gall
18	44	42.7	330	1 PE42_ARATH	O95b81 arabidopsis
19	44	42.7	496	1 SGK3_HUMAN	O96b71 homo sapien
20	44	42.7	539	1 RCCL_SCHPO	P28745 schizosacch
21	43	41.7	170	1 CRAA_BRAVA	O02487 bradypus va
22	43	41.7	171	1 IPP1_MOUSE	O99rt9 mus musculu
23	43	41.7	287	1 UPK3_MOUSE	O99jx8 mus musculu
24	43	41.7	460	1 VL2_HPV44	O80918 human papil
25	43	41.7	496	1 SGK3_MOUSE	O99e13 mus musculu
26	43	41.7	559	1 TRAP_PLAFA	P16893 plasmodium
27	43	41.7	687	1 SRB4_YEAST	P32569 saccharomyc
28	43	41.7	713	1 CADD_HUMAN	P55290 homo sapien
29	43	41.7	714	1 CADD_MOUSE	O9wtr5 mus musculu
30	43	41.7	2223	1 CCRAE_DISOM	P56699 discopysge o
31	42.5	41.3	1078	1 CARR_BUCAI	P57244 buchnera ap
32	42	40.8	130	1 CLP1_DROME	P02839 drosophila
33	42	40.8	173	1 CRAA_CHICK	P02504 gallus gall

34	42	40.8	173	1 CRAA_KHEAM	P02505 rhea amerie
35	42	40.8	464	1 CDK8_HUMAN	P49336 homo sapien
36	42	40.8	522	1 ABP1_SCHPO	P49777 schizosacch
37	42	40.8	674	1 YFE2_YEAST	P45660 saccharomyc
38	42	40.8	1182	1 HAIR_MOUSE	O61645 mus musculu
39	42	40.8	1461	1 TOP2_CANAL	P87078 candida alb
40	41.5	40.3	4289	1 TENX_HUMAN	P22105 homo sapien
41	41	39.8	116	1 SELX_MOUSE	O91lc3 mus musculu
42	41	39.8	123	1 TX12_MOUSE	O99r81 mus musculu
43	41	39.8	173	1 CRAA_TUPTTE	P02506 tupinambis
44	41	39.8	268	1 ESL1_MYCGE	O49412 mycoplasma
45	41	39.8	295	1 YUST_BACSU	O32186 bacillus su

ALIGNMENTS

RESULT 1

ID	SGK1_HUMAN	STANDARD:	PRT:	431 AA.
AC	O00141: O9UN56:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)			
DE	(Serum/glucocorticoid-regulated kinase 1).			
GN	SGK OR SGK1			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97272242; PubMed=9114008;			
RA	Waldegger S., Barth P., Raber G., Lang F.;			
RT	"Cloning and characterization of a putative human serine/threonine			
RT	protein kinase transcriptionally modified during anisotonic and			
RT	isotonic alterations of cell volume."			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:4440-4445(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98390195; PubMed=9722955;			
RA	Waldegger S., Erdel M., Nagl U.O., Barth P., Raber G., Steuer S.,			
RA	Utermann G., Paulmichl M., Lang F.;			
RT	"Genomic organization and chromosomal localization of the human SGK			
RT	protein kinase gene."			
RL	Genomics 51:299-302(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Dermal papilla;			
RA	Kim M.K., Kim Y.H., Seo J.M., Lee H.M., Chung H.J., Sohn M.Y.,			
RA	Hwang S.Y., Im S.U., Jung E.J., Lee J.H., Kim J.C.;			
RT	"A catalogue of genes in the human dermal papilla cells as identified			
RT	by expressed sequence tags."			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cervix;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fabey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smallick D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-20018032; PubMed-10548550;
 RA Kobayashi T., Deak M., Morrice N., Cohen P.,
 RT "Characterization of the structure and regulation of two novel
 RT isoforms of serum- and glucocorticoid-induced protein kinase.";
 RL Biochem. J. 344:189-197(1999).
 RN [6]
 RP PHOSPHORYLATION ON THR-256, AND MUTAGENESIS OF THR-256 AND SER-422.
 RX MEDLINE-99208518; PubMed-10191262;
 RA Kobayashi T., Cohen P.,
 RT "Activation of serum- and glucocorticoid-regulated protein kinase by
 RT agonists that activate phosphatidylinositol 3-kinase is mediated by
 RT 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2.";
 RL Biochem. J. 339:319-328(1999).
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE-20345128; PubMed-10884438;
 RA Lang F., Klingel K., Wagner C.A., Stegen C., Waerntges S.,
 RA Friedrich B., Lanzendoerfer M., Melzig J., Moschen I., Steuer S.,
 RA Waldeger S., Sauter M., Paulmichl M., Gerke V., Risler T., Gamba G.,
 RA Capasso G., Kandolf R., Hebert S.C., Massry S.G., Broer S.,
 RT "Deranged transcriptional regulation of cell-volume-sensitive kinase
 RT hSKG in diabetic nephropathy.",
 KL Proc. Natl. Acad. Sci. U.S.A. 97:8157-8162(2000).
 RN [8]
 RP FUNCTION.
 RX MEDLINE-20584913; PubMed-11154281;
 RA Brunet A., Park J., Tran H., Hu L.S., Hemmings B.A., Greenberg M.E.,
 RT "Protein kinase SGK mediates survival signals by phosphorylating the
 RT forward transcription factor FKHL1 (FOXO3a).",
 RL Mol. Cell. Biol. 21:952-963(2001).
 RN [9]
 RP FUNCTION.
 RX MEDLINE-22284526; PubMed-12397388;
 RA Gamber N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
 RA Huber S.M., Kobayashi T., Cohen P., Lang F.,
 RT "K(+) channel activation by all three isoforms of serum- and
 RT glucocorticoid-dependent protein kinase SGK.",
 RL Pflugers Arch. 445:60-66(2002).
 CC -!- FUNCTION: Protein kinase that plays an important role in cellular
 CC stress response. Activates certain potassium, sodium, and chloride
 CC channels, suggesting an involvement in the regulation of processes
 CC such as cell survival, neuronal excitability, and renal sodium
 CC excretion. Sustained high levels and activity may contribute to
 CC conditions such as hypertension and diabetic nephropathy. Mediates
 CC cell survival signals, phosphorylates and negatively regulates
 CC pro-apoptotic FOXO3a.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 CC phosphorylation.
 CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 CC in the pancreas, followed by placenta, kidney and lung.
 CC -!- INDUCTION: By serum and/or glucocorticoids. By excessive
 CC extracellular glucose and by TGF-beta, in cultured cells.
 CC -!- PTM: Regulated by phosphorylation. Phosphoinositide 3-kinase (PI3-
 CC kinase) pathway promotes phosphorylation at Ser-422 which in turn
 CC increases the phosphorylation of Thr-256 by PDK1.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y10032; CA71138.1; -
 DR EMBL; AJ000512; CA04146.1; -
 DR EMBL; AF153609; AAD41091.1; -
 DR EMBL; BC001263; AAH01263.1; -
 DR HSP; P00517; IYDR.
 DR Genew; HGNC:10810; SGK.
 DR MIM; 602958; -
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO; GO:0006950; P:response to stress; TAS.
 DR GO; GO:0006814; P:sodium ion transport; TAS.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase_C; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW Apoptosis; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 98 355 PROTEIN KINASE
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 256 256 PHOSPHORYLATION (BY PDK1).
 FT DOMAIN 131 141 LYS/GLU-RICH.
 FT MUTAGEN 256 256 T->A: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->D: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->E: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->A: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->D: 10-FOLD ACTIVATION.
 FT CONFLICT 381 381 E -> D (IN REF. 3 AND 4).
 SQ SEQUENCE 431 AA; 48956 MW; F3697DA5707339D CRC64;
 Query Match 100.0%; Score 103; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DPEFTEPVPNSIGKSPDS 19
 Db 386 DPEFTEPVPNSIGKSPDS 404
 RESULT 2
 ID SGK1_RAT STANDARD; PRT; 430 AA.
 AC Q06226;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
 DE (Serum/glucocorticoid-regulated kinase 1).
 OS SGK OR SGK1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer 344;
 RX MEDLINE-93204949; PubMed-8455596;
 RA Webster M.K., Goya T., Ge F., Malyar A.C., Firestone G.L.,
 RT "Characterization of sgk, a novel member of the serine/threonine
 RT protein kinase gene family which is transcriptionally induced by
 RT glucocorticoids and serum.",
 RL Mol. Cell. Biol. 13:2031-2040(1993).
 RN [2]

RP INDUCTION BY CNS INJURY.
RX MEDLINE-95157173; PubMed-7854047;
RA Imaizumi K., Tsuda M., Watanaka A., Tohyama M., Takagi T.;
RT "Differential expression of sgk mRNA, a member of the Ser/Thr protein
RT kinase gene family, in rat brain after CNS injury.";
RL Brain Res. Mol. Brain Res. 26:189-196(1994).
RN [3]
RP INDUCTION BY FSH.
RX MEDLINE-95258633; PubMed-7740159;
RA Richards J.S., Fitzpatrick S.L., Clemens J.W., Morris J.K.,
RA Alliston T., Sirote J.;
RT "Ovarian cell differentiation: a cascade of multiple hormones,
RT cellular signals, and regulated genes.";
RL Recent Prog. Horm. Res. 50:223-254(1995).
RN [4]
RP INDUCTION BY P53.
RX TISSUE=Mammary epithelium;
RX MEDLINE-96218163; PubMed-8647846;
RA Maiyar A.C., Huang A.J., Phu P.T., Cha H.H., Firestone G.L.;
RT "p53 stimulates promoter activity of the sgk.
RT serum/glucocorticoid-inducible serine/threonine protein kinase gene
RT in rodent mammary epithelial cells.";
RL J. Biol. Chem. 271:12414-12422(1996).
RN [5]
RP PHOSPHORYLATION ON THR-256 BY PDPK1.
RX MEDLINE-99286226; PubMed-10357815;
RA Park J., Leong M.L., Buse P., Maiyar A.C., Firestone G.L.,
RA Hemmings B.A.;
RT "Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI
RT 3-kinase-stimulated signaling pathway.";
RL EMBO J. 18:3024-3033(1999).
RN [6]
RP FUNCTION: Protein kinase that plays an important role in
CC activating certain potassium, sodium, and chloride channels,
CC cell survival, neuronal excitability, and renal sodium excretion.
CC May be a key component of cellular stress response (By
CC similarity). May also play an important role in the development of
CC particular groups of neurons in the postnatal brain.
CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
CC phosphorylation (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
CC in the ovary, thymus and lung.
CC -!- INDUCTION: By dexamethasone and serum. By tumor suppressor p53 in
CC mammary epithelial tumor cells. By FSH in granulosa cells. By
CC injury to the central nervous system.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L01624; AAA42137.1;
CC HSPSP; P00517; 1YDR.
CC InterPro: IPR000961; pkinase.C.
CC InterPro: IPR000719; prot_kinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase.1.
CC Pfam: PF00433; pkinase.C.1.
CC ProDom: PD000001; Prot_kinase.1.
CC SMART: SM00133; S_TK_X.1.
CC SMART: SM00220; S_TK_X.1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
CC Apoptosis: Transferrase: Serine/threonine-protein kinase: ATP-binding;
CC Phosphorylation.
CC DOMAIN 98 354 PROTEIN KINASE.
CC NP_BIND 104 112 ATP (BY SIMILARITY).
CC

FT BINDING 127 127 ATP (BY SIMILARITY).
FT ACT_SITE 222 222 BY SIMILARITY.
FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1).
SQ SEQUENCE 430 AA; 48927 MW; 0D5645B04156F26D CRC64;
Query Match 92.2%; Score 95; DB 1; Length 430;
Best Local Similarity 89.5%; Pred. No. 2.9e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 DPEFTEEPVNSIGKSPUS 19
Db 385 DPEFTEEPVPSIGRSPDS 403
|||||
RESULT 3
SGK_MOUSE STANDARD; PRT; 431 AA.
ID SGK_MOUSE AC Q9WVC6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
DE (Serum/glucocorticoid-regulated kinase 1).
GN SGK OR SGK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99287894; PubMed-10358046;
RA Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
RA Fejes-Toth G.;
RT "Sgk is an aldosterone-induced kinase in the renal collecting duct.
RT Effects on epithelial Na⁺ channels.";
RL J. Biol. Chem. 274:16973-16978(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20215285; PubMed-10751222;
RA Shigaev A., Asher C., Latter H., Garty H., Keuveny E.;
RT "Regulation of sgk by aldosterone and its effects on the epithelial
RT Na(+) channel.";
RL Am. J. Physiol. 278:F613-F619(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=FVB/N;
RX MEDLINE-22388257; PubMed-12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,
RA Villalón D., Muzny E., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP REGULATION BY P53.
RC TISSUE=Mammary epithelium;
RX MEDLINE-96218163; PubMed-8647846;
RA Maiyar A.C., Huang A.J., Phu P.T., Cha H.H., Firestone G.L.;
RT "p53 stimulates promoter activity of the sgk.

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.1)
 DE (Serum/glucocorticoid-regulated kinase 2).
 GN Sgk2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20018032; PubMed=10548550;
 RA Kobayashi T., Deak M., Morrice N., Cohen P.;
 RT "Characterization of the structure and regulation of two novel
 RT isoforms of serum- and glucocorticoid-induced protein kinase.";
 RL Biochem. J. 344:189-197(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=22354883; PubMed=12468851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H.A., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustingich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.O., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelain A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Vektard R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Vasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.H., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION: Involved in the activation of potassium channels (By

CC similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9QZS5-1; Sequence-Displayed;
 CC Name=2;
 CC IsoId=Q9QZS5-2; Sequence-VSP_004933;
 CC Note=NO experimental confirmation available;
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- CAUTION: Not regulated by serum or glucocorticoids.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF159033; AAC12756.1;
 CC EMBL: AK050003; BAC34031.1;
 CC EMBL: EC026549; AAC26549.1;
 CC HSP: P05132; ICTP.
 CC MGD: MGI:1351318; Sgk2.
 CC InterPro: IPR000961; Pkinase.C.
 CC InterPro: IPR000719; Prot_Kinase.
 CC InterPro: IPR002290; Ser_thr_kinase.
 CC Pfam: PF00069; pkinase; 1.
 CC Pfam: PF00433; pkinase.C; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00133; S_TK_X; 1.
 CC SMART: SM00220; S_TK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC Transferrin: Serine/threonine-protein kinase: ATP-binding;
 CC Phosphorylation; Alternative splicing.
 CC DOMAIN 35 292 PROTEIN_KINASE.
 CC NP_BIND 41 49 ATP (BY SIMILARITY).
 CC BINDING 64 64 ATP (BY SIMILARITY).
 CC ACT_SITE 159 159 BY SIMILARITY.
 CC MOD_RES 193 193 PHOSPHORYLATION (BY PDPK1)
 CC (BY SIMILARITY).
 CC VARSPLIT 171 199 Missing (in isoform 2).
 CC CONFLICT 77 77 MISSING (IN REF. 3).
 CC SEQUENCE 367 AA; 41359 MW; 688C04B1A1E9E33A CRC64;
 Query Match 62.1%; Score 64; DB 1; Length 367;
 Best Local Similarity 63.2%; Pred. No. 0.016;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 1 DPEETEEVPVNSICKSPDS 19
 DB 323 DPEFTQEAIVSKSIGCTPDT 341
 RESULT 6
 Sgk2_HUMAN
 ID Sgk2_HUMAN STANDARD; PRT; 427 AA.
 AC Q9HBY8; Q9UKG6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.37)
 DE (Serum/glucocorticoid-regulated kinase 2).
 GN Sgk2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE-21638749; PubMed-11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Brown A.J.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaesthai M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.L.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Stewart C.A., Sulston J.E.,
RA Swann R.M., Symamore N., Taylor R., Taylor M., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.N.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT The DNA sequence and comparative analysis of human chromosome 20.;
RL Nature 414:865-871(2001).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE-Colon;
RX MEDLINE-22388257; PubMed-12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.S.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN FUNCTION.
RX MEDLINE-22284526; PubMed-12397388;
RA Gampert N., Fillon Y., Feng Y., Friedrich B., Lang P.A., Henke G.,
RA Huber S.M., Kobayashi T., Cohen P., Lang P.;
RT "K(+)" channel activation by all three isoforms of serum- and
RT glucocorticoid-dependent protein kinase SGK";
RL Pflugers Arch. 445:60-66(2002).
CC -!- FUNCTION: Involved in the activation of potassium channels.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms-2;

CC Name-2: Synonyms-beta;
CC IsoId-Q9HBY8-1; Sequence-Displayed;
CC Name-1: Synonyms-alpha;
CC IsoId-Q9HBY8-2; Sequence-VSP_004932;
CC -!- TISSUE SPECIFICITY: Highly expressed in liver, kidney and
CC pancreas, and at lower levels in brain.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- CAUTION: Not regulated by serum or glucocorticoids.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF186470; AAG17012.1; -
CC EMBL; AF169034; AAF12757.2; -
CC EMBL; 298752; CAC18509.1; -
CC EMBL; BC014037; AAL14037.1; -
CC HSSP; FG5132; IATP.
CC Gene; HGNC:13900; SGK2.
CC MIM; 607583; -
CC InterPro; IPR000961; Pkinase_C.
CC InterPro; IPR000719; Prot_kinase.
CC Pfam; PF00469; Pkinase_1.
CC Pfam; PF00433; Pkinase_C; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 95 352 PROTEIN KINASE.
FT NP_BIND 101 109 ATP (BY SIMILARITY).
FT BINDING 124 124 ATP (BY SIMILARITY).
FT ACT_SITE 219 219 HY SIMILARITY.
FT MOD_RES 253 253 PHOSPHORYLATION (BY POPK1).
FT VARSPPLIC 1 60 Missing (in isoform 1).
FT MUTAGEN 416 416 S->D: INCREASED ACTIVATION.
SQ SEQUENCE 427 AA; 47604 MW; D8F0FA6DF54B1370 CRC64;
Query Match 62.1%; Score 64; DB 1; Length 427;
Best Local Similarity 63.2%; Pred. No. 0.019;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 DPEETEEVPVNSICKSPUS 19
DB 383 DPEETEEVPVNSICKSPUS 401
RESULT 7
ID VL2_HPV55 STANDARD; PRT; 460 AA.
AC Q80939;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Minor capsid protein L2.
GN L2.
OS Human papillomavirus type 55.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=37114;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: U31791; AAA79483.1; ..
 DR InterPro: IPR000784; Late_L2.
 DR Pfam: PF00513; late_protein.
 KW Coat protein; late protein.
 SQ SEQUENCE 460 AA; 49339 MW; 7B9BD6F0E94E00AE CRC64;

 Query Match 56.3%; Score 58; DB 1; Length 460;
 Best Local Similarity 68.8%; Pred. No. 0.18;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

 QY 1 DPFTPEPVPNSTGKS 16
 II:IIIIII I I
 Db 356 DPFTPEPVPNSTGKS 371

 RESULT 8
 ID EXTN_TOBAC STANDARD; PRT; 620 AA.
 AC P13983;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).
 GN HRPNT3.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotianaceae;
 OC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Xanthi; TISSUE=Leaf;
 RA MEDLINE=90128263; PubMed=2612909;
 RX Keller B., Lamb C.J.;
 RT "Specific expression of a novel cell wall hydroxyproline-rich
 RT glycoprotein gene in lateral root initiation."
 RL Genes Dev. 3:1639-1646(1989).
 CC -!- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN
 CC THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE
 CC MAIN ROOT.
 CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -!- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
 CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
 CC GLYCOSYLATED.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: X13885; CAA32090.1; ..
 DR PIR; S06733; S06733.
 DR InterPro: IPR000480; Glutelin.
 DR InterPro: IPR002965; P-rich_extensin.
 DR PRINTS; PR00211; GLUTELIN.
 DR PRINTS; PR01217; PRICHEXTENSIN.
 KW Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
 KW Hydroxylation.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 620 EXTENSIN.
 FT REPEAT 70 73 H-A-P-P.

 FT REPEAT 148 151 H-A-P-P.
 FT DOMAIN 229 242 2 X 7 AA TANDEM REPEATS OF T-H-K-H-A-P-P.
 FT REPEAT 229 235 1.
 FT REPEAT 236 242 2.
 FT DOMAIN 205 620 CONTAINS THE SER-PRO(4) REPEATS.
 FT DOMAIN 499 600 3 X APPROXIMATE TANDEM REPEATS.
 SQ SEQUENCE 620 AA; 65406 MW; 641DD2278AB28524 CRC64;

 Query Match 49.5%; Score 51; DB 1; Length 620;
 Best Local Similarity 55.8%; Pred. No. 3;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

 QY 2 PEFTPEPVPNSIGKSPDS 19
 I I I I I I I I I I
 Db 36 PPVTSQPPSSIGLSPPS 53

 RESULT 9
 ID UPK3_HUMAN STANDARD; PRT; 287 AA.
 AC Q75631; Q60261;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Uroplakin III precursor (UPIII).
 GN UPK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ureter;
 RA Gaall K., Hall G., Smith B., Southgate J.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bladder urothelium;
 RX MEDLINE=99035381; PubMed=9818021;
 RA Yuasa T., Yoshiki T., Tanaka T., Kim C.J., Isono T., Okada Y.;
 RT "Expression of uroplakin Ib and uroplakin III genes in tissues and
 RT peripheral blood of patients with transitional cell carcinoma."
 RL Jpn. J. Cancer Res. 89:879-882(1998).
 RN [3]
 RP SEQUENCE OF 33-287 FROM N.A.
 RC TISSUE=Urinary bladder;
 RA Yuasa T., Tanaka T., Isono T.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: COMPONENT OF THE ASYMMETRIC UNIT MEMBRANE (AUM); A
 CC HIGHLY SPECIALIZED BIOMEMBRANE ELABORATED BY TERMINALLY
 CC DIFFERENTIATED UROTHELIAL CELLS. MAY PLAY AN IMPORTANT ROLE
 CC IN AUM-CYTOSKELETON INTERACTION IN TERMINALLY DIFFERENTIATED
 CC UROTHELIAL CELLS. IT ALSO CONTRIBUTES TO THE FORMATION OF
 CC PREVENTING BACTERIAL ADHERENCE (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL: AF085808; AAC34888.1; ..
 DR EMBL; AB010637; BAA31460.1; ..
 DR EMBL; AB010116; BAA25678.1; ..
 DR Genew; HGNC:12580; UPK3.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT CHAIN 19 287 UROPLAKIN III.
 FT DOMAIN 19 207 LUMENAL (POTENTIAL).

FT TRANSMEM 208 235 POTENTIAL.
 FT DOMAIN 236 287 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 154 154 P -> A (IN REF. 1).
 FT CONFLICT 236 236 A -> D (IN REF. 1).
 SQ SEQUENCE 287 AA; 30651 MW; DFL452048E4C58F CRC64;
 Query Match 45.6%; Score 47; DB 1; Length 287;
 Best Local Similarity 47.4%; Pred. No. 5.3;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 1 DPEETPEVPNSIGKSPDS 19
 I : | | | | | | | | | |
 Db 247 DSQITQBAVPSLGLGASESS 265

RESULT 10
 ID SLPI_CLOTH STANDARD; PRT; 1664 AA.
 AC Q08852;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Cell surface glycoprotein 1 precursor (Outer layer protein B) (S-layer protein 1).
 DE protein 1).
 GN OLEP.
 OS Clostridium thermoecellum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID-1515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCIE 10682;
 RX MEDLINE-93209931; PubMed-8458632;
 RA Fujino T., Beguin P., Aubert J.-P.;
 RT "Organization of a Clostridium thermoecellum gene cluster encoding the cellulosomal scaffolding protein CIPa and a protein possibly involved in attachment of the cellulosome to the cell surface."
 RL J. Bacteriol. 175:1891-1899(1993).
 CC -!- SUBUNIT: ASSSEMBLED INTO MONO-LAYERED CRYSTALLINE ARRAYS.
 CC -!- SUBCELLULAR LOCATION: Cell wall.
 CC -!- SIMILARITY: Contains 4 S-layer homology (SLH) domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 DR EMBL; X67506; CAA47841.1; .
 DR PIR; T18262; T18262.
 DR InterPro; IPR001119; SLH.
 DR Pfam; PF00395; SLH; 3.
 DR PROSITE; PS01072; SLH_DOMAIN; 2.
 KW Cell wall; S-layer; Signal; Repeat.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 1664 CELL SURFACE GLYCOPROTEIN 1.
 FT DOMAIN 36 763 4 X 156 AA APPROXIMATE REPEATS.
 FT REPEAT 36 191 1.
 FT REPEAT 207 363 2.
 FT REPEAT 409 565 3.
 FT REPEAT 607 763 4.
 FT DOMAIN 771 1377 APPROXIMATE TANDEM REPEATS OF 3-P-S-D-E-P.
 FT DOMAIN 1378 1449 GLY/PRO/SER/THR-RICH.
 FT DOMAIN 1453 1494 SLH 1 (INCOMPLETE).
 FT DOMAIN 1495 1565 SLH 2.
 FT DOMAIN 1566 1625 SLH 3.
 FT DOMAIN 1626 1646 SLH 4 (INCOMPLETE).

SQ SEQUENCE 1664 AA; 178194 MW; 5F396695BA9FE74B CRC64;
 Query Match 45.1%; Score 46.5; DB 1; Length 1664;
 Best Local Similarity 55.6%; Pred. No. 47;
 Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Oy 2 PEETPEVPNSIGKSPDS 19
 I | | | | | | | | | |
 Db 391 PTVTEPVPSEL---PDS 405

RESULT 11
 ID TNFA_CANEA STANDARD; PRT; 233 AA.
 AC PS1742; Q28339;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tumor necrosis factor precursor (TNF-alpha) (Tumor necrosis factor ligand superfamily member 2) (TNF-a) (cachectin).
 GN TNF OR TNFSF2 OR TNFA.
 OS Canis familiaris (Dog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID-9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fiers W.;
 RT "Tumour necrosis factor.";
 RL (In) Sim E. (eds.);
 RL The natural immune system humoral factors, pp.65-119, IRL Press, Oxford (1993).
 RN [12]
 RP SEQUENCE FROM N.A.
 RA Zucker K., Lu P., Fuller L., Asthana D., Esquenazi V., Miller J.;
 RT "Cloning and expression of the cDNA for canine tumor necrosis factor-alpha in E. coli";
 RL Lymphokine Res. 13:191-196(1994).
 RN [3]
 RP SEQUENCE OF 74-205 FROM N.A.
 RC STRAIN-Beagle; TISSUE-Blood;
 RA Gilmore W.H., Carter S.D., Bennett M., Barnes A., Kelly D.F.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Cytokine that binds to TNFRSF1A/TNFR1 and TNFRSF1B/TNFR. It is mainly secreted by macrophages and can induce cell death of certain tumor cell lines. It is potent pyrogen causing fever by direct action or by stimulation of interleukin 1 secretion and is implicated in the induction of cachexia. Under certain conditions it can stimulate cell proliferation and induce cell differentiation.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).
 CC -!- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).
 CC -!- PTM: The membrane form, but not the soluble form, is phosphorylated on serine residues. Dephosphorylation of the membrane form occurs by binding to soluble TNFRSF1A/TNFR1 (By similarity).
 CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION.
 CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 DR EMBL; X94932; CAA64403.1; .

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
RN (4)
RP SEQUENCE OF 295-464 FROM N.A.
RC STRAIN=CD16-1;
RX MEDLINE=21270454; PubMed=11376151;
RA Watanabe T., Miyashita K., Saito T.T., Yoneki T., Kakiyama Y.,
RA Nabeshima K., Kishi Y.A., Shimoda C., Nojima H.;
RT "Comprehensive isolation of meiosis-specific genes identifies novel
RT proteins and unusual non-coding transcripts in Schizosaccharomyces
RT pombe";
RL Nucleic Acids Res. 29:2327-2337(2001).
CC -!- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U29891; AAB53688.2; -
DR EMBL; 298601; CAB11273.1; -
DR EMBL; AB054312; BAB60879.1; -
DR GeneDB_Spombe; SPAC24C9.15C; -
DR InterPro: IPR000038; GTP_Cell_Div.
DR Pfam: PF00735; GTP_COC; 1.
DR ProDom: PD002565; GTP_Cell_Div; 1.
KW GTP-binding; Coiled coil; Meiosis.
FT NP-BIND 125 132 GTP (POTENTIAL).
FT DOMAIN 396 453 COILED COIL (POTENTIAL).
SQ SEQUENCE 464 AA; 53104 MW; F7BE83DE827ECE1B CRC64;
Query Match 44.7%; Score 46; DB 1; Length 464;
Best Local Similarity 64.3%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 EFTEEPVPSNGKS 16
II :: |||||
Db 76 EFKQDEVPSNGKS 89
RESULT 14
YLES_CABEL STANDARD; PRT; 837 AA.
AC P46941;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C38D4.5 in chromosome III.
GN C38D4.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Coles L.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC -!- SIMILARITY: Contains 1 WW domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 246241; CAA86318.1; -
DR PIR; T19825; T19825.
DR WormPep; C38D4.5; CE00918.
DR InterPro: IPR001849; PH.
DR InterPro: IPR00198; RhoGAP.
DR InterPro: IPR001202; WW_Rsp5_WMP.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00620; RhoGAP; 1.
DR Pfam: PF00397; WW; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00324; RHO-GAP; 1.
DR SMART; SM00456; WW; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50238; RHO-GAP; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS50020; WW_DOMAIN_2; 1.
KW Hypothetical protein.
FT DOMAIN 96 129 WW.
FT DOMAIN 386 505 PH.
FT DOMAIN 610 827 RHO-GAP.
SQ SEQUENCE 837 AA; 94196 MW; D1895622D1F5997 CRC64;
Query Match 44.7%; Score 46; DB 1; Length 837;
Best Local Similarity 56.2%; Pred. No. 26;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 DPEETPEVPNSIGKS 16
II :: |||||
Db 238 DSDFDDEPVTSSRKA 253
RESULT 15
BMPA_BORGA STANDARD; PRT; 325 AA.
AC Q31357; Q31360;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Basic membrane protein A precursor (Immunodominant antigen P39)
DE (Fragment).
GN BMPA.
OS Borrelia garinii.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=29519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PBI, and PLI;
RX MEDLINE=98010210; PubMed=9350727;
RA Roessler D., Hauser U., Wilske B.;
RT "Heterogeneity of BmpA (P39) among European isolates of Borrelia
RT burgdorferi sensu lato and influence of interspecies variability on
RT serodiagnosis";
RL J. Clin. Microbiol. 35:2752-2758(1997).
CC -!- FUNCTION: NOT KNOWN; IMMUNOGENIC PROTEIN.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (probable).
CC -!- SIMILARITY: BELONGS TO THE BMP FAMILY OF LIPOPROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X97244; CAA65883.1; -

DR EMBL: X97238; CAA65877.1: -.
DR InterPro: IPR003760; Bmp.
DR InterPro: IPR000437; Prok_lipoprot.
DR Pfam: PF02608; Bmp; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; PARTIAL.
KW Antigen; Membrane; Lipoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 3
FT CHAIN 4 325
FT LIPID 4 4
FT VARIANT 8 8
FT VARIANT 12 12
FT VARIANT 93 93
FT VARIANT 111 111
FT VARIANT 180 180
FT VARIANT 205 205
FT VARIANT 239 239
FT VARIANT 243 243
FT VARIANT 250 250
FT VARIANT 257 257
SQ SEQUENCE 325 AA; 35480 MW; D19281A2A4B1C158 CRC64;
PROBABLE.
BASIC MEMBRANE PROTEIN A.
N-ACYL DIGLYCERIDE (PROBABLE).
G -> D (IN STRAIN PL1).
S -> N (IN STRAIN PL1).
P -> S (IN STRAIN PL1).
T -> A (IN STRAIN PL1).
N -> D (IN STRAIN PL1).
S -> G (IN STRAIN PL1).
I -> V (IN STRAIN PL1).
A -> S (IN STRAIN PL1).
S -> A (IN STRAIN PL1).
N -> S (IN STRAIN PL1).

Query Match 43.7%; Score 45; DB 1; Length 325;
Best Local Similarity 46.2%; Pred. No. 13;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DPETTEEPVPSI 13
DB 101 DPVISEEPIPTNL 113
||::|||:|::|

Search completed: August 11, 2003, 08:19:54
Job time : 3.87554 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: August 11, 2003, 08:05:45 ; Search time 2.5279 Seconds
(without alignments)
722.816 Million cell updates/sec

Title: US-10-000-039a-3

Perfect score: 103

Sequence: 1 DPEFTEPVPNSIGKSPDS 19

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	95	92.2	431	2 A48094	serum and glucocor
2	51	49.5	620	2 S06733	hydroxyproline-ric
3	48	46.6	448	2 T48966	hypothetical prote
4	48	46.6	490	2 A92993	mannitol dehydroge
5	48	46.6	490	2 A92990	mannitol dehydroge
6	47.5	46.1	163	2 H84488	hypothetical prote
7	47.5	46.1	4006	2 T09070	probable tenascin
8	47	45.6	490	2 G71563	hypothetical prote
9	46.5	45.1	1664	2 T18262	S-layer protein -
10	46	44.7	287	2 T45986	uroplakin iii - bo
11	46	44.7	464	2 T38356	sepin homolog spn
12	46	44.7	708	2 A92626	ferrichrome iron r
13	46	44.7	747	2 G97408	fesA protein U6140
14	46	44.7	837	2 T18825	hypothetical prote
15	45	43.7	339	2 A84496	probable replicati
16	45	43.7	374	2 T50265	probable phosphate
17	45	43.7	548	2 T32686	hypothetical prote
18	44	42.7	144	2 T01290	hypothetical prote
19	44	42.7	323	2 T05478	peroxidase (EC 1.1
20	44	42.7	481	2 T10036	hypothetical prote
21	44	42.7	539	2 T50368	piml GTPase protei
22	44	42.7	539	2 B40039	piml hypothetical A
23	44	42.7	640	2 S62747	homeotic protein A
24	44	42.7	1111	2 G96693	hypothetical prote
25	43.5	42.2	1089	2 T31583	hypothetical prote
26	43	41.7	162	2 F84380	hypothetical prote
27	43	41.7	170	1 CY0W43	alpha-crystallin c
28	43	41.7	188	2 AE1233	Type-I signal pept
29	43	41.7	274	2 T05509	hypothetical prote

ALIGNMENTS

RESULT 1

A48094 serum and glucocorticoid-regulated kinase - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 19-Dec-1997

C:Accession: A48094

R:Webster, M.K.; Goya, L.; Ge, Y.; Maiyar, A.C.; Firestone, G.L.

Mol. Cell. Biol. 13, 2031-2040, 1993

A:Title: Characterization of sgk, a novel member of the serine/threonine protein kinase

A:Reference number: A48094; MUID:93204949; PMID:8455596

A:Accession: A48094

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-431 <WEB>

A:Experimental source: Con8, hdb mammary epithelial tumor cells

A:Note: sequence extracted from NCBI backbone (NCBI:127618, NCBI:127619)

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

C:Keywords: ATP

F:96-385/Domain: protein kinase homology <KIN>

F:104-112/Region: protein kinase ATP-binding motif

Query Match 92.2%; Score 95; DB 2; Length 431;

Best Local Similarity 89.5%; Pred. No. 3.3e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DPEFTEPVPNSIGKSPDS 19

Db 386 DPEFTEPVPSSIGKSPDS 404

RESULT 2

S06733

hydroxyproline-rich glycoprotein precursor - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jul-2000

C:Accession: S06733

R:Keller, B.; Lamb, C.J.

Genes Dev. 3, 1639-1646, 1989

A:Title: Specific expression of a novel cell wall hydroxyproline-rich glycoprotein ge

A:Reference number: S06733; MUID:90128263; PMID:2612909

A:Accession: S06733

A:Molecule type: DNA

A:Residues: 1-620 <KEI>

A:Cross-references: EMBL:X13885; NID:q19866; PID:CAA32090.1; PID:q19867

C:Superfamily: hydroxyproline-rich glycoprotein

C:Keywords: glycoprotein

Query Match 49.5%; Score 51; DB 2; Length 620;

Best Local Similarity 55.6%; Pred. No. 5.4;

Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 PEFTEPVPNSIGKSPDS 19

```

Db      36 PPVTSQPPSSIGLSPPS 53

RESULT 3
T48966
hypothetical protein F4M19.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48966
R:Vitale, D.; Liguori, R.; Flores, M.; Argiriou, A.; De Simone, V.; Mewes, H.W.; Rudd, S.
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25010
A:Accession: T48966
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <VIT>
A:Cross-references: EMBL:AL356013; GSPDB:GN00063; ATSP:F4M19.70
A:Experimental source: cultivar Columbia; BAC clone F4M19
C:Genetics:
A:Gene: ATSP:F4M19.70
A:Map position: 3
A:Introns: 47/3; 80/3; 93/2; 135/1; 163/2; 206/3; 271/2; 312/3; 383/3; 411/1

Query Match 46.6%; Score 48; DB 2; Length 448;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 DPFTTEPVPNSIGSKS 16
DB      315 DPENLPKPKNLIGKT 330

RESULT 4
A99293
mannitol dehydrogenase PA2342 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Dec-2002
C:Accession: A99293
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A99293
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK99867.1; PID:gl5159810; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L2604
A:Map position: linear chromosome
C:Superfamily: mannitol 2-dehydrogenase

Query Match 46.6%; Score 48; DB 2; Length 490;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      1 DPFTTEPVPNSIGSKS 19
DB      423 DPFAEAWVPTGNGEPPDA 441

RESULT 5
AH2990
mannitol dehydrogenase uxub [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Dec-2002
C:Accession: AH2990
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perly, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH2990
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <KUR>
A:Cross-references: GB:AE008689; PIDN:AL44342.1; PID:gl7741934; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: uxuB
A:Map position: linear chromosome
C:Superfamily: mannitol 2-dehydrogenase

Query Match 46.6%; Score 48; DB 2; Length 490;
Best Local Similarity 47.4%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY      1 DPFTTEPVPNSIGSKS 19
DB      423 DPFAEAWVPTGNGEPPDA 441

RESULT 6
H84488
hypothetical protein At2g09840 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: H84488
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.B.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84488
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-163 <STO>
A:Cross-references: GB:AE002093; NID:g4803918; PIDN:AAD9792.1; GSPDB:GN00179
C:Genetics:
A:Gene: At2g09840
A:Map position: 2
C:Superfamily: Arabidopsis hypothetical protein T20F6.2

Query Match 46.1%; Score 47.5; DB 2; Length 163;
Best Local Similarity 45.8%; Pred. No. 3.9;
Matches 11; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

QY      1 DPFTTEPVPNSIGSKS 19
DB      61 DPDEDEPPVSPVPEALHSPDS 84

RESULT 7
T09070
probable tenascin X - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000
C:Accession: T09070
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahern, M.E.; Dankers, C.; Lasky, S.; Loretz, C.;
submitted to the EMBL Data Library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: Z16543
A:Accession: T09070
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4006 <ROW>
A:Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564958
C:Genetics:
A:Gene: TNX
A:Map position: 17

```

A: Introns: 124/1; 735/1; 773/3; 826/1; 914/1; 1037/1; 1135/1; 1232/1; 1329/1; 1440/1; 1501/1; 3115/1; 3208/1; 3302/1; 3405/1; 3517/1; 3558/1; 3606/1; 3694/1; 3737/3; 3789/1; 3826/1; 3894/1; 3962/1; 4030/1; 4098/1; 4166/1; 4234/1; 4302/1; 4370/1; 4438/1; 4506/1; 4574/1; 4642/1; 4710/1; 4778/1; 4846/1; 4914/1; 4982/1; 5050/1; 5118/1; 5186/1; 5254/1; 5322/1; 5390/1; 5458/1; 5526/1; 5594/1; 5662/1; 5730/1; 5798/1; 5866/1; 5934/1; 6002/1; 6070/1; 6138/1; 6206/1; 6274/1; 6342/1; 6410/1; 6478/1; 6546/1; 6614/1; 6682/1; 6750/1; 6818/1; 6886/1; 6954/1; 7022/1; 7090/1; 7158/1; 7226/1; 7294/1; 7362/1; 7430/1; 7498/1; 7566/1; 7634/1; 7702/1; 7770/1; 7838/1; 7906/1; 7974/1; 8042/1; 8110/1; 8178/1; 8246/1; 8314/1; 8382/1; 8450/1; 8518/1; 8586/1; 8654/1; 8722/1; 8790/1; 8858/1; 8926/1; 8994/1; 9062/1; 9130/1; 9198/1; 9266/1; 9334/1; 9402/1; 9470/1; 9538/1; 9606/1; 9674/1; 9742/1; 9810/1; 9878/1; 9946/1; 10014/1; 10082/1; 10150/1; 10218/1; 10286/1; 10354/1; 10422/1; 10490/1; 10558/1; 10626/1; 10694/1; 10762/1; 10830/1; 10898/1; 10966/1; 11034/1; 11102/1; 11170/1; 11238/1; 11306/1; 11374/1; 11442/1; 11510/1; 11578/1; 11646/1; 11714/1; 11782/1; 11850/1; 11918/1; 11986/1; 12054/1; 12122/1; 12190/1; 12258/1; 12326/1; 12394/1; 12462/1; 12530/1; 12598/1; 12666/1; 12734/1; 12802/1; 12870/1; 12938/1; 13006/1; 13074/1; 13142/1; 13210/1; 13278/1; 13346/1; 13414/1; 13482/1; 13550/1; 13618/1; 13686/1; 13754/1; 13822/1; 13890/1; 13958/1; 14026/1; 14094/1; 14162/1; 14230/1; 14298/1; 14366/1; 14434/1; 14502/1; 14570/1; 14638/1; 14706/1; 14774/1; 14842/1; 14910/1; 14978/1; 15046/1; 15114/1; 15182/1; 15250/1; 15318/1; 15386/1; 15454/1; 15522/1; 15590/1; 15658/1; 15726/1; 15794/1; 15862/1; 15930/1; 15998/1; 16066/1; 16134/1; 16202/1; 16270/1; 16338/1; 16406/1; 16474/1; 16542/1; 16610/1; 16678/1; 16746/1; 16814/1; 16882/1; 16950/1; 17018/1; 17086/1; 17154/1; 17222/1; 17290/1; 17358/1; 17426/1; 17494/1; 17562/1; 17630/1; 17698/1; 17766/1; 17834/1; 17902/1; 17970/1; 18038/1; 18106/1; 18174/1; 18242/1; 18310/1; 18378/1; 18446/1; 18514/1; 18582/1; 18650/1; 18718/1; 18786/1; 18854/1; 18922/1; 18990/1; 19058/1; 19126/1; 19194/1; 19262/1; 19330/1; 19398/1; 19466/1; 19534/1; 19602/1; 19670/1; 19738/1; 19806/1; 19874/1; 19942/1; 20010/1; 20078/1; 20146/1; 20214/1; 20282/1; 20350/1; 20418/1; 20486/1; 20554/1; 20622/1; 20690/1; 20758/1; 20826/1; 20894/1; 20962/1; 21030/1; 21098/1; 21166/1; 21234/1; 21302/1; 21370/1; 21438/1; 21506/1; 21574/1; 21642/1; 21710/1; 21778/1; 21846/1; 21914/1; 21982/1; 22050/1; 22118/1; 22186/1; 22254/1; 22322/1; 22390/1; 22458/1; 22526/1; 22594/1; 22662/1; 22730/1; 22798/1; 22866/1; 22934/1; 23002/1; 23070/1; 23138/1; 23206/1; 23274/1; 23342/1; 23410/1; 23478/1; 23546/1; 23614/1; 23682/1; 23750/1; 23818/1; 23886/1; 23954/1; 24022/1; 24090/1; 24158/1; 24226/1; 24294/1; 24362/1; 24430/1; 24498/1; 24566/1; 24634/1; 24702/1; 24770/1; 24838/1; 24906/1; 24974/1; 25042/1; 25110/1; 25178/1; 25246/1; 25314/1; 25382/1; 25450/1; 25518/1; 25586/1; 25654/1; 25722/1; 25790/1; 25858/1; 25926/1; 25994/1; 26062/1; 26130/1; 26198/1; 26266/1; 26334/1; 26402/1; 26470/1; 26538/1; 26606/1; 26674/1; 26742/1; 26810/1; 26878/1; 26946/1; 27014/1; 27082/1; 27150/1; 27218/1; 27286/1; 27354/1; 27422/1; 27490/1; 27558/1; 27626/1; 27694/1; 27762/1; 27830/1; 27898/1; 27966/1; 28034/1; 28102/1; 28170/1; 28238/1; 28306/1; 28374/1; 28442/1; 28510/1; 28578/1; 28646/1; 28714/1; 28782/1; 28850/1; 28918/1; 28986/1; 29054/1; 29122/1; 29190/1; 29258/1; 29326/1; 29394/1; 29462/1; 29530/1; 29598/1; 29666/1; 29734/1; 29802/1; 29870/1; 29938/1; 30006/1; 30074/1; 30142/1; 30210/1; 30278/1; 30346/1; 30414/1; 30482/1; 30550/1; 30618/1; 30686/1; 30754/1; 30822/1; 30890/1; 30958/1; 31026/1; 31094/1; 31162/1; 31230/1; 31298/1; 31366/1; 31434/1; 31502/1; 31570/1; 31638/1; 31706/1; 31774/1; 31842/1; 31910/1; 31978/1; 32046/1; 32114/1; 32182/1; 32250/1; 32318/1; 32386/1; 32454/1; 32522/1; 32590/1; 32658/1; 32726/1; 32794/1; 32862/1; 32930/1; 32998/1; 33066/1; 33134/1; 33202/1; 33270/1; 33338/1; 33406/1; 33474/1; 33542/1; 33610/1; 33678/1; 33746/1; 33814/1; 33882/1; 33950/1; 34018/1; 34086/1; 34154/1; 34222/1; 34290/1; 34358/1; 34426/1; 34494/1; 34562/1; 34630/1; 34698/1; 34766/1; 34834/1; 34902/1; 34970/1; 35038/1; 35106/1; 35174/1; 35242/1; 35310/1; 35378/1; 35446/1; 35514/1; 35582/1; 35650/1; 35718/1; 35786/1; 35854/1; 35922/1; 35990/1; 36058/1; 36126/1; 36194/1; 36262/1; 36330/1; 36398/1; 36466/1; 36534/1; 36602/1; 36670/1; 36738/1; 36806/1; 36874/1; 36942/1; 37010/1; 37078/1; 37146/1; 37214/1; 37282/1; 37350/1; 37418/1; 37486/1; 37554/1; 37622/1; 37690/1; 37758/1; 37826/1; 37894/1; 37962/1; 38030/1; 38098/1; 38166/1; 38234/1; 38302/1; 38370/1; 38438/1; 38506/1; 38574/1; 38642/1; 38710/1; 38778/1; 38846/1; 38914/1; 38982/1; 39050/1; 39118/1; 39186/1; 39254/1; 39322/1; 39390/1; 39458/1; 39526/1; 39594/1; 39662/1; 39730/1; 39798/1; 39866/1; 39934/1; 40002/1; 40070/1; 40138/1; 40206/1; 40274/1; 40342/1; 40410/1; 40478/1; 40546/1; 40614/1; 40682/1; 40750/1; 40818/1; 40886/1; 40954/1; 41022/1; 41090/1; 41158/1; 41226/1; 41294/1; 41362/1; 41430/1; 41498/1; 41566/1; 41634/1; 41702/1; 41770/1; 41838/1; 41906/1; 41974/1; 42042/1; 42110/1; 42178/1; 42246/1; 42314/1; 42382/1; 42450/1; 42518/1; 42586/1; 42654/1; 42722/1; 42790/1; 42858/1; 42926/1; 42994/1; 43062/1; 43130/1; 43198/1; 43266/1; 43334/1; 43402/1; 43470/1; 43538/1; 43606/1; 43674/1; 43742/1; 43810/1; 43878/1; 43946/1; 44014/1; 44082/1; 44150/1; 44218/1; 44286/1; 44354/1; 44422/1; 44490/1; 44558/1; 44626/1; 44694/1; 44762/1; 44830/1; 44898/1; 44966/1; 45034/1; 45102/1; 45170/1; 45238/1; 45306/1; 45374/1; 45442/1; 45510/1; 45578/1; 45646/1; 45714/1; 45782/1; 45850/1; 45918/1; 45986/1; 46054/1; 46122/1; 46190/1; 46258/1; 46326/1; 46394/1; 46462/1; 46530/1; 46598/1; 46666/1; 46734/1; 46802/1; 46870/1; 46938/1; 47006/1; 47074/1; 47142/1; 47210/1; 47278/1; 47346/1; 47414/1; 47482/1; 47550/1; 47618/1; 47686/1; 47754/1; 47822/1; 47890/1; 47958/1; 48026/1; 48094/1; 48162/1; 48230/1; 48298/1; 48366/1; 48434/1; 48502/1; 48570/1; 48638/1; 48706/1; 48774/1; 48842/1; 48910/1; 48978/1; 49046/1; 49114/1; 49182/1; 49250/1; 49318/1; 49386/1; 49454/1; 49522/1; 49590/1; 49658/1; 49726/1; 49794/1; 49862/1; 49930/1; 49998/1; 50066/1; 50134/1; 50202/1; 50270/1; 50338/1; 50406/1; 50474/1; 50542/1; 50610/1; 50678/1; 50746/1; 50814/1; 50882/1; 50950/1; 51018/1; 51086/1; 51154/1; 51222/1; 51290/1; 51358/1; 51426/1; 51494/1; 51562/1; 51630/1; 51698/1; 51766/1; 51834/1; 51902/1; 51970/1; 52038/1; 52106/1; 52174/1; 52242/1; 52310/1; 52378/1; 52446/1; 52514/1; 52582/1; 52650/1; 52718/1; 52786/1; 52854/1; 52922/1; 52990/1; 53058/1; 53126/1; 53194/1; 53262/1; 53330/1; 53398/1; 53466/1; 53534/1; 53602/1; 53670/1; 53738/1; 53806/1; 53874/1; 53942/1; 54010/1; 54078/1; 54146/1; 54214/1; 54282/1; 54350/1; 54418/1; 54486/1; 54554/1; 54622/1; 54690/1; 54758/1; 54826/1; 54894/1; 54962/1; 55030/1; 55098/1; 55166/1; 55234/1; 55302/1; 55370/1; 55438/1; 55506/1; 55574/1; 55642/1; 55710/1; 55778/1; 55846/1; 55914/1; 55982/1; 56050/1; 56118/1; 56186/1; 56254/1; 56322/1; 56390/1; 56458/1; 56526/1; 56594/1; 56662/1; 56730/1; 56798/1; 56866/1; 56934/1; 57002/1; 57070/1; 57138/1; 57206/1; 57274/1; 57342/1; 57410/1; 57478/1; 57546/1; 57614/1; 57682/1; 57750/1; 57818/1; 57886/1; 57954/1; 58022/1; 58090/1; 58158/1; 58226/1; 58294/1; 58362/1; 58430/1; 58498/1; 58566/1; 58634/1; 58702/1; 58770/1; 58838/1; 58906/1; 58974/1; 59042/1; 59110/1; 59178/1; 59246/1; 59314/1; 59382/1; 59450/1; 59518/1; 59586/1; 59654/1; 59722/1; 59790/1; 59858/1; 59926/1; 59994/1; 60062/1; 60130/1; 60198/1; 60266/1; 60334/1; 60402/1; 60470/1; 60538/1; 60606/1; 60674/1; 60742/1; 60810/1; 60878/1; 60946/1; 61014/1; 61082/1; 61150/1; 61218/1; 61286/1; 61354/1; 61422/1; 61490/1; 61558/1; 61626/1; 61694/1; 61762/1; 61830/1; 61898/1; 61966/1; 62034/1; 62102/1; 62170/1; 62238/1; 62306/1; 62374/1; 62442/1; 62510/1; 62578/1; 62646/1; 62714/1; 62782/1; 62850/1; 62918/1; 62986/1; 63054/1; 63122/1; 63190/1; 63258/1; 63326/1; 63394/1; 63462/1; 63530/1; 63598/1; 63666/1; 63734/1; 63802/1; 63870/1; 63938/1; 64006/1; 64074/1; 64142/1; 64210/1; 64278/1; 64346/1; 64414/1; 64482/1; 64550/1; 64618/1; 64686/1; 64754/1; 64822/1; 64890/1; 64958/1; 65026/1; 65094/1; 65162/1; 65230/1; 65298/1; 65366/1; 65434/1; 65502/1; 65570/1; 65638/1; 65706/1; 65774/1; 65842/1; 65910/1; 65978/1; 66046/1; 66114/1; 66182/1; 66250/1; 66318/1; 66386/1; 66454/1; 66522/1; 66590/1; 66658/1; 66726/1; 66794/1; 66862/1; 66930/1; 66998/1; 67066/1; 67134/1; 67202/1; 67270/1; 67338/1; 67406/1; 67474/1; 67542/1; 67610/1; 67678/1; 67746/1; 67814/1; 67882/1; 67950/1; 68018/1; 68086/1; 68154/1; 68222/1; 68290/1; 68358/1; 68426/1; 68494/1; 68562/1; 68630/1; 68698/1; 68766/1; 68834/1; 68902/1; 68970/1; 69038/1; 69106/1; 69174/1; 69242/1; 69310/1; 69378/1; 69446/1; 69514/1; 69582/1; 69650/1; 69718/1; 69786/1; 69854/1; 69922/1; 69990/1; 70058/1; 70126/1; 70194/1; 70262/1; 70330/1; 70398/1; 70466/1; 70534/1; 70602/1; 70670/1; 70738/1; 70806/1; 70874/1; 70942/1; 71010/1; 71078/1; 71146/1; 71214/1; 71282/1; 71350/1; 71418/1; 71486/1; 71554/1; 71622/1; 71690/1; 71758/1; 71826/1; 71894/1; 71962/1; 72030/1; 72098/1; 72166/1; 72234/1; 72302/1; 72370/1; 72438/1; 72506/1; 72574/1; 72642/1; 72710/1; 72778/1; 72846/1; 72914/1; 72982/1; 73050/1; 73118/1; 73186/1; 73254/1; 73322/1; 73390/1; 73458/1; 73526/1; 73594/1; 73662/1; 73730/1; 73798/1; 73866/1; 73934/1; 74002/1; 74070/1; 74138/1; 74206/1; 74274/1; 74342/1; 74410/1; 74478/1; 74546/1; 74614/1; 74682/1; 74750/1; 74818/1; 74886/1; 74954/1; 75022/1; 75090/1; 75158/1; 75226/1; 75294/1; 75362/1; 75430/1; 75498/1; 75566/1; 75634/1; 75702/1; 75770/1; 75838/1; 75906/1; 75974/1; 76042/1; 76110/1; 76178/1; 76246/1; 76314/1; 76382/1; 76450/1; 76518/1; 76586/1; 76654/1; 76722/1; 76790/1; 76858/1; 76926/1; 76994/1; 77062/1; 77130/1; 77198/1; 77266/1; 77334/1; 77402/1; 77470/1; 77538/1; 77606/1; 77674/1; 77742/1; 77810/1; 77878/1; 77946/1; 78014/1; 78082/1; 78150/1; 78218/1; 78286/1; 78354/1; 78422/1; 78490/1; 78558/1; 78626/1; 78694/1; 78762/1; 78830/1; 78898/1; 78966/1; 79034/1; 79102/1; 79170/1; 79238/1; 79306/1; 79374/1; 79442/1; 79510/1; 79578/1; 79646/1; 79714/1; 79782/1; 79850/1; 79918/1; 79986/1; 80054/1; 80122/1; 80190/1; 80258/1; 80326/1; 80394/1; 80462/1; 80530/1; 80598/1; 80666/1; 80734/1; 80802/1; 80870/1; 80938/1; 81006/1; 81074/1; 81142/1; 81210/1; 81278/1; 81346/1; 81414/1; 81482/1; 81550/1; 81618/1; 81686/1; 81754/1; 81822/1; 81890/1; 81958/1; 82026/1; 82094/1; 82162/1; 82230/1; 82298/1; 82366/1; 82434/1; 82502/1; 82570/1; 82638/1; 82706/1; 82774/1; 82842/1; 82910/1; 82978/1; 83046/1; 83114/1; 83182/1; 83250/1; 83318/1; 83386/1; 83454/1; 83522/1; 83590/1; 83658/1; 83726/1; 83794/1; 83862/1; 83930/1; 83998/1; 84066/1; 84134/1; 84202/1; 84270/1; 84338/1; 84406/1; 84474/1; 84542/1; 84610/1; 84678/1; 84746/1; 84814/1; 84882/1; 84950/1; 85018/1; 85086/1; 85154/1; 85222/1; 85290/1; 85358/1; 85426/1; 85494/1; 85562/1; 85630/1; 85698/1; 85766/1; 85834/1; 85902/1; 85970/1; 86038/1; 86106/1; 86174/1; 86242/1; 86310/1; 86378/1; 86446/1; 86514/1; 86582/1; 86650/1; 86718/1; 86786/1; 86854/1; 86922/1; 86990/1; 87058/1; 87126/1; 87194/1; 87262/1; 87330/1; 87398/1; 87466/1; 87534/1; 87602/1; 87670/1; 87738/1; 87806/1; 87874/1; 87942/1; 88010/1; 88078/1; 88146/1; 88214/1; 88282/1; 88350/1; 88418/1; 88486/1; 88554/1; 88622/1; 88690/1; 88758/1; 88826/1; 88894/1; 88962/1; 89030/1; 89098/1; 89166/1; 89234/1; 89302/1; 89370/1; 89438/1; 89506/1; 89574/1; 89642/1; 89710/1; 89778/1; 89846/1; 89914/1; 89982/1; 90050/1; 90118/1; 90186/1; 90254/1; 90322/1; 90390/1; 90458/1; 90526/1; 90594/1; 90662/1; 90730/1; 90798/1; 90866/1; 90934/1; 91002/1; 91070/1; 91138/1; 91206/1; 91274/1; 91342/1; 91410/1; 91478/1; 91546/1; 91614/1; 91682/1; 91750/1; 91818/1; 91886/1; 91954/1; 92022/1; 92090/1; 92158/1; 92226/1; 92294/1; 92362/1; 92430/1; 92498/1; 92566/1; 92634/1; 92702/1; 92770/1; 92838/1; 92906/1; 92974/1; 93042/1; 93110/1; 93178/1; 93246/1; 93314/1; 93382/1; 93450/1; 93518/1; 93586/1; 93654/1; 93722/1; 93790/1; 93858/1; 93926/1; 93994/1; 94062/1; 94130/1; 94198/1; 94266/1; 94334/1; 94402/1; 94470/1; 94538/1; 94606/1; 94674/1; 94742/1; 94810/1; 94878/1; 94946/1; 95014/1; 95082/1; 95150/1; 95218/1; 95286/1; 95354/1; 95422/1; 95490/1; 95558/1; 9

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:08:14 ; Search time 6.03433 Seconds
(without alignments)
499.774 Million cell updates/sec

Title: US-10-000-039a-3

Perfect score: 103

Sequence: 1 DPEFTEPVPNSICKSPDS 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

Rank	ID	Score	Query Match	Length	DB ID	Description
1	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	103	100.0	19	AAW77218	Human cell-volume
2	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	103	100.0	321	AAW99836	ACC protein kinase
3	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	103	100.0	373	AAW99815	ACC protein kinase
4	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	103	100.0	388	AAW48573	Human breast tumou
5	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	103	100.0	431	AAW77217	Human cell-volume
6	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	103	100.0	431	AAW54025	Human protein kinase
7	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	103	100.0	431	AAW90139	Human sgk protein.
8	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	103	100.0	431	AAW24115	Human serum and gl
9	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	103	100.0	431	AAW95279	Human serum and gl
10	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*					
11	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*					
12	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*					
13	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*					
14	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*					
15	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*					
16	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*					
17	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*					
18	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*					
19	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*					
20	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*					
21	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*					
22	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*					
23	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*					
24	/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*					

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	103	100.0	19	AAW77218	Human cell-volume
2	103	100.0	321	AAW99836	ACC protein kinase
3	103	100.0	373	AAW99815	ACC protein kinase
4	103	100.0	388	AAW48573	Human breast tumou
5	103	100.0	431	AAW77217	Human cell-volume
6	103	100.0	431	AAW54025	Human protein kinase
7	103	100.0	431	AAW90139	Human sgk protein.
8	103	100.0	431	AAW24115	Human serum and gl
9	103	100.0	431	AAW95279	Human serum and gl

10	103	100.0	431	22	AAW65613	Novel protein kinase
11	95	92.2	430	22	AAW65614	Novel protein kinase
12	95	92.2	431	21	AAW24116	Rat serum and gl
13	95	92.2	431	21	AAW93530	A rat serum and gl
14	64	62.1	244	22	AAW65615	Novel protein kinase
15	64	62.1	319	22	AAW99838	ACC protein kinase
16	64	62.1	367	21	AAW95275	Human serum and gl
17	64	62.1	367	21	AAW95277	Mouse serum and gl
18	64	62.1	367	22	AAU28087	Novel human secret
19	64	62.1	367	22	AAW65708	Novel protein kinase
20	64	62.1	367	23	AAE23765	Human serum and gl
21	64	62.1	382	22	AAW25594	Human protein sequ
22	64	62.1	427	21	AAW95276	Human serum and gl
23	47	45.6	287	23	ABG61848	Prostate cancer as
24	47	45.6	287	24	ABR48223	Human bladder can
25	47	45.6	496	20	AAW37175	Amino acid sequen
26	46.5	45.1	1664	19	AAW43106	C. thermocellum ol
27	46	44.7	93	21	AAW57060	Human prostate can
28	46	44.7	142	22	ABR95928	Human testicular a
29	46	44.7	142	22	AAW95231	Human reproductive
30	46	44.7	403	23	ABP73244	Candida albicans e
31	45.5	44.2	1245	23	ABR77587	Human cancer assoc
32	45.5	44.2	1491	23	AAE25344	Human patched-like
33	45.5	44.2	1492	23	ABR77588	Human cancer assoc
34	45	43.7	256	22	AAU14724	Novel bone marrow
35	45	43.7	1175	22	ABG07962	Novel human diagno
36	45	43.7	1706	22	AAW79270	Human protein sequ
37	44	42.7	170	21	AAW09728	Arabidopsis thalia
38	44	42.7	298	21	AAW09727	Arabidopsis thalia
39	44	42.7	299	21	AAW09726	Arabidopsis thalia
40	44	42.7	322	22	AAW99837	ACC protein kinase
41	44	42.7	323	21	AAW07631	Arabidopsis thalia
42	44	42.7	323	21	AAW36807	Arabidopsis thalia
43	44	42.7	323	23	ABR93042	Herbicidally activ
44	44	42.7	324	21	AAW07630	Arabidopsis thalia
45	44	42.7	324	21	AAW36806	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW77218
ID AAW77218 standard; Peptide: 19 AA.
XX AC AAW77218;
XX AC AAW77218;
XX DF 16-NOV-1998 (first entry)
XX DE Human cell-volume regulating kinase h-syk peptide fragment.
XX DE Human: cell-volume; Kinase; h-syk; diabetes mellitus;
KW renal insufficiency; inflammation; Alzheimer's disease.
XX OS Homo sapiens.
XX PN EP861896-A2.
XX PD 02-SEP-1998.
XX PF 27-JAN-1998; 98EP-0101338.
XX PR 28-FEB-1997; 97DE-1008173.
XX (DADE-) DADE BEHRING MABURG GMBH.
XX Lang F, Waldegger S;
XX WPI; 1998-449109/39.
XX New nucleic acid encoding cell-volume regulating kinase h-syk and
PT related proteins - used for diagnosis and treatment of diseases
PT involving changes in cell volume, e.g. renal insufficiency,

PT inflammation, infections etc.
 XX
 PS Claim 7; Page 9; 15pp; German.
 XX
 CC The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.
 XX
 SQ Sequence 19 AA;
 Query Match 100.0%; Score 103; DB 19; Length 19;
 Best Local Similarity 100.0%; Pred. No. 4.2e-09;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DPEFTEPVPNSIGKSPDS 19
 |||||
 DB 1 DPEFTEPVPNSIGKSPDS 19
 |||||
 RESULT 2
 AAB99836
 JD AAB99836 standard; Protein; 321 AA.
 XX
 AC AAB99836;
 XX
 DT 20-SEP-2001 (first entry)
 XX
 DE AGC protein kinase family member SGK1 protein sequence.
 XX
 KW Protein kinase; identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN W0200144497-A2.
 XX
 PD 21-JUN-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04598.
 XX
 PR 02-DEC-1999; 99US-0168559.
 XX
 PA (UYDU-) UNIV DUNDEE.
 XX
 PI Alessi D, Biondi R;
 XX
 DR WPI; 2001-390252/41.
 XX
 PT Identifying modulators of protein kinase (PK) activity, useful in
 PT developing drugs for treating cancer or diabetes, by measuring the
 PT ability of the compound to modulate or mimic the interaction of PK with
 PT interacting polypeptides -
 XX
 PS Disclosure; Fig 16; 180pp; English.
 XX
 CC The present invention describes a method for identifying a compound that
 CC modulates protein kinase activity. The method comprises measuring the
 CC ability of the compound to inhibit, promote or mimic the interaction of
 CC a hydrophobic pocket-containing protein kinase with an interacting
 CC polypeptide. The interacting polypeptide interacts with the hydrophobic
 CC pocket of the protein kinase and/or comprises the amino acid sequence
 CC Phe/Tyr-Xaa-Xaa-Phe/Tyr (I). The method is useful in screening assays
 CC for developing pharmaceutical compounds or drugs. Compounds, polypeptides
 CC or polynucleotides from the present invention are useful in medicine,
 CC in need of modulation of signalling by a hydrophobic pocket-containing

CC particularly in the manufacture of a medicament for treating a patient
 CC in need of modulation of signalling by a hydrophobic pocket-containing
 CC protein kinase. Specifically, the patient has cancer or diabetes or is
 CC in need of inhibition of apoptosis, e.g. a patient suffering from tissue
 CC injury or ischaemic injury, including stroke. The compound or
 CC composition is also useful for inhibiting the degree or rate of
 CC phosphorylation by the protein kinase. The interacting polypeptide or
 CC compound is useful in methods of stabilising a hydrophobic pocket-
 CC containing protein kinase, where the protein kinase is exposed to the
 CC compound or polypeptide. AAB99836 to AAB99847 represent amino acid
 CC sequences, and AAB44210 and AAB44211 represent oligonucleotide sequences,
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 321 AA;
 Query Match 100.0%; Score 103; DB 22; Length 321;
 Best Local Similarity 100.0%; Pred. No. 9.6e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DPEFTEPVPNSIGKSPDS 19
 |||||
 DB 276 DPEFTEPVPNSIGKSPDS 294
 |||||
 RESULT 3
 AAB99815
 ID AAB99815 standard; Protein; 373 AA.
 XX
 AC AAB99815;
 XX
 DT 20-SEP-2001 (first entry)
 XX
 DE AGC protein kinase family member SGK protein sequence.
 XX
 KW Protein kinase; identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN W0200144497-A2.
 XX
 PD 21-JUN-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04598.
 XX
 PR 02-DEC-1999; 99US-0168559.
 XX
 PA (UYDU-) UNIV DUNDEE.
 XX
 PI Alessi D, Biondi R;
 XX
 DR WPI; 2001-390252/41.
 XX
 PT Identifying modulators of protein kinase (PK) activity, useful in
 PT developing drugs for treating cancer or diabetes, by measuring the
 PT ability of the compound to modulate or mimic the interaction of PK with
 PT interacting polypeptides -
 XX
 PS Disclosure; Fig 15; 180pp; English.
 XX
 CC The present invention describes a method for identifying a compound that
 CC modulates protein kinase activity. The method comprises measuring the
 CC ability of the compound to inhibit, promote or mimic the interaction of
 CC a hydrophobic pocket-containing protein kinase with an interacting
 CC polypeptide. The interacting polypeptide interacts with the hydrophobic
 CC pocket of the protein kinase and/or comprises the amino acid sequence
 CC Phe/Tyr-Xaa-Xaa-Phe/Tyr (I). The method is useful in screening assays
 CC for developing pharmaceutical compounds or drugs. Compounds, polypeptides
 CC or polynucleotides from the present invention are useful in medicine,
 CC particularly in the manufacture of a medicament for treating a patient
 CC in need of modulation of signalling by a hydrophobic pocket-containing

CC protein kinase. Specifically, the patient has cancer or diabetes or is
 CC in need of inhibition of apoptosis, e.g. a patient suffering from tissue
 CC injury or ischaemic injury, including stroke. The compound or
 CC composition is also useful for inhibiting the degree or rate of
 CC phosphorylation by the protein kinase. The interacting polypeptide or
 CC compound is useful in methods of stabilising a hydrophobic pocket.
 CC containing protein kinase, where the protein kinase is exposed to the
 CC compound or polypeptide. AAB99786 to AAB99847 represent amino acid
 CC sequences, and AAH44210 and AAH44211 represent oligonucleotide sequences,
 CC used in the exemplification of the present invention.

XX Sequence 373 AA;

Query Match 100.0%; Score 103; DB 22; Length 373;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNSIGKSPDS 19

DB 328 DPEETPEVPNSIGKSPDS 346

RESULT 4

AA48573

ID AAY48573 standard; Protein: 388 AA.

XX AAY48573;

XX AAY48573;

XX 08-DEC-1999 (first entry)

XX Human breast tumour-associated protein 34.

XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;

XX treatment; tumour; cytostatic; medicament.

XX Homo sapiens.

XX DE19813839-AL.

XX 23-SEP-1999.

XX 20-MAR-1998; 98DE-1013839.

XX 20-MAR-1998; 98DE-1013839.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosentahl A;

XX WPI: 1999-528981/45.

XX N-PSDB; AA233647.

XX Human nucleic acid sequences and protein products from tumor breast

XX tissue, useful for breast cancer therapy.

XX Claim 22; 157-158; 188pp; German.

XX This invention describes novel human nucleic acid sequences from tumor

XX breast tissue which have cytostatic activity. The nucleic acid sequences

XX can be used to produce and isolate full-length gene sequences. They can

XX be used to express proteins, which can be used as tools to find an

XX activity against breast cancer. The sequences can be used in sense or

XX antisense form. They are especially useful for medicaments for gene

XX therapy to treat breast cancer. AAY48540-Y48617 represent protein

XX fragments encoded by the expressed sequence tags described in the method

XX of the invention.

XX Sequence 388 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 103; DB 20; Length 388;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNSIGKSPDS 19

DB 343 DPEETPEVPNSIGKSPDS 363

RESULT 5

AAW77217

ID AAW77217 standard; Protein: 431 AA.

XX AAW77217;

XX AAW77217;

XX 16-NOV-1998 (first entry)

XX Human cell-volume regulating kinase h-sgk.

XX Human; cell-volume; kinase; h-sgk; diabetes mellitus;

XX renal insufficiency; inflammation; Alzheimer's disease.

XX Homo sapiens.

XX EP861896-A2.

XX 02-SEP-1998.

XX 27-JAN-1998; 98EP-0101338.

XX 28-FEB-1997; 97DE-1008173.

XX (DADE-) DADE BEHRING MARBURG GMBH.

XX Lang F, Waldegger S;

XX WPI: 1998-449109/39.

XX N-PSDB; AAY48311.

XX New nucleic acid encoding cell-volume regulating kinase h-sgk and

XX related proteins - used for diagnosis and treatment of diseases

XX involving changes in cell volume, e.g. renal insufficiency,

XX inflammation, infections etc.

XX Disclosure; Fig 1; 15pp; German.

XX The human cell-volume regulating kinase h-sgk is inhibited by the

XX swelling of cells (or presence of urea), whereas cell shrinkage

XX stimulates its expression. The nucleic acid h-sgk, and fragments, are

XX particularly used to detect changes in cell volume, specifically for

XX diagnosis of conditions that involve such changes, e.g. hyper- and hypo-

XX natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,

XX hepatic encephalopathy, inflammation, microbial/viral infection, fructose

XX intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.

XX The nucleic acid, protein and products including receptors that bind

XX h-sgk, can be used to treat these disorders.

XX Sequence 431 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 103; DB 19; Length 431;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNSIGKSPDS 19

DB 386 DPEETPEVPNSIGKSPDS 404

RESULT 6

AAW54025

ID AAW54025 standard; Protein: 431 AA.

XX AAW54025;

XX AAW54025;

XX 31-JUL-1998 (first entry)

XX Human protein kinase HPK-3.

XX Protein kinase; human; HPK; signalling cascade; kinase expression;
 KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;
 KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;
 KW lymphoma; therapy.
 XX Homo sapiens.
 XX WO9811234-A2.
 XX 19-MAR-1998.
 XX 10-SEP-1997; 97WO-US15923.
 XX 12-SEP-1996; 96US-0712709.
 XX (INCY-) INCYTE PHARM INC.
 XX Au-Young J, Guegler KJ, Hawkins PR;
 PI WPI: 1998-207394/18.
 DR N-PSDB; AAV23833.
 XX New isolated human protein kinase(s) - used to develop products for
 PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
 PT asthma, multiple sclerosis or rheumatoid arthritis
 XX Claim 23; Fig 3; 75pp; English.
 XX This sequence is a human protein kinase (HPK) of the invention. The
 CC HPK protein can be used to develop products for studying signalling
 CC cascades in various cells and tissues, diagnosing disease and selecting
 CC inhibitors or drugs with the potential to intervene in various disorders
 CC or diseases in which altered kinase expression is implicated. The
 CC products can be used to e.g. reverse memory loss such as due to
 CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
 CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
 CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
 XX lymphocytic leukaemias or lymphomas.
 XX Sequence 431 AA;
 Query Match 100.0%; Score 103; DB 19; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DPEFTEPVPNSIGKSPDS 19
 DB 386 DPEFTEPVPNSIGKSPDS 404
 RESULT 7
 AAW90139
 ID AAW90139 standard; Protein; 431 AA.
 XX AC AAW90139;
 XX 15-MAR-1999 (first entry)
 DT Human sgk protein.
 DE Serum glucocorticoid regulated kinase; sgk; human; treatment; inhibitor;
 XX serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 KW chronic renal failure; inflammation; Alzheimers disease; wound.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 381
 FT /note= "encoded by GAG"
 XX EP887081-A2.

PD 30-DEC-1998.
 XX 27-MAY-1998; 98EP-0304189.
 XX 27-JUN-1997; 97US-0051124.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Kumar JM;
 XX WPI: 1999-047627/05.
 DR N-PSDB; AAV74190.
 XX Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 PT disease - by administration of nucleic acids and antagonists which
 PT inhibit activity or expression of human serum glucocorticoid
 PT regulated kinase (sgk), a serine/threonine protein kinase
 XX Disclosure; Page 16-17; 17pp; English.
 XX This sequence represents a novel human serum glucocorticoid regulated
 CC kinase (sgk) protein which is a member of the serine/threonine protein
 CC kinase family. This protein is used for the treatment of a subject having
 CC need to inhibit/antagonise activity or expression of human sgk
 CC polypeptide e.g. for the treatment of chronic renal failure, diabetic
 CC nephropathy, inflammation, Alzheimers disease and wounds.
 XX Sequence 431 AA;
 Query Match 100.0%; Score 103; DB 20; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DPEFTEPVPNSIGKSPDS 19
 DB 386 DPEFTEPVPNSIGKSPDS 404
 RESULT 8
 AAB24115
 ID AAB24115 standard; Protein; 431 AA.
 XX AC AAB24115;
 XX 29-JAN-2001 (first entry)
 DT Human serum and glucocorticoid regulated kinase protein (HSGK).
 XX Human; STK3; serine threonine kinase; HSGK; RSGK;
 XX serum and glucocorticoid regulated kinase.
 XX Homo sapiens.
 XX CN1259573-A.
 XX 12-JUL-2000.
 XX 29-OCT-1998; 98CN-0123822.
 XX 29-OCT-1998; 98CN-0123822.
 XX (UYFU-) UNIV FUDAN.
 XX Yu L, Fu Q, Zhao Y;
 XX WPI: 2000-587991/56.
 XX New human serine threonine protein kinase, its code sequence,
 PT preparation, and use -
 XX Example 3; Fig 3; 29pp; Chinese.
 XX The present invention describes human serine threonine kinase STK3,
 CC

CC which is a new member of the serine/threonine protein kinase family.
 CC Human SGK shares homology with human serum and glucocorticoid
 CC regulated kinase (HSGK) and rat SGK (RSGK). The present sequence
 CC represents the human HSGK protein sequence from the present invention.

XX SQ Sequence 431 AA;
 Query Match 100.0%; Score 103; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNSIGKSPDS 19
 ID AAY95279 standard; Protein; 431 AA.
 DB 386 DPEETPEVPNSIGKSPDS 404

RESULT 9
 AAY95279
 ID AAY95279 standard; Protein; 431 AA.
 AC AAY95279;

XX 12-SEP-2000 (first entry)
 DT Human serum and glucocorticoid-induced protein kinase.

DE Human serum and glucocorticoid-induced protein kinase.

XX Serum and glucocorticoid-induced protein kinase; SGK;
 KW human; phosphorylation; cancer; diabetes; ischaemia; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers
 FJI Modified-site 256
 FT /note= "O-phosphorylated"
 FT Modified-site 422
 FT /note= "O-phosphorylated"

XX WO200035946-A1.

XX 22-JUN-2000.

XX 14-DEC-1999; 99WO-GB04232.

XX 14-DEC-1998; 98US-0112217.

XX 19-AUG-1999; 99GB-0019676.

XX (UYDU-) UNIV DUNDEE.

XX Cohen P, Kobayashi T, Deak M;

XX WPI; 2000-442364/38.

XX Modulation of serum and glucocorticoid-induced protein kinase (SGK) by
 PT phosphorylation with 3-phosphoinositide-dependent protein kinase-1
 PT (PDK1) or dephosphorylation, useful for treatment of cancer, diabetes
 PT and ischemic diseases.

XX Disclosure; Fig 13; 127pp; English.

XX The present sequence is that of human serum and
 CC glucocorticoid-induced protein kinase (SGK). SGK was initially
 CC identified as a glucocorticoid and osmotic stress-responsive gene.
 CC It is activated by phosphorylation of its Thr-308 and Ser-472
 CC residues. Novel isoforms of SGK, i.e. SGK2 and SGK3 (see
 CC AAY95275-78), have been identified. The invention provides methods
 CC of activating SGK activity by phosphorylation using
 CC 3-phosphoinositide-dependent protein kinase-1 (PDK1), and of
 CC reducing the activity of SGK by dephosphorylation. The invention
 CC also provides a method of identifying a compound that modulates the
 CC activity of SGK. Such compounds are useful for treating patients
 CC requiring modulation of SGK, such as patients with cancer, diabetes
 CC or ischaemic disease.

SQ Sequence 431 AA;

Query Match 100.0%; Score 103; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPEETPEVPNSIGKSPDS 19
 ID AAB65613 standard; Protein; 431 AA.
 DB 386 DPEETPEVPNSIGKSPDS 404

RESULT 10
 AAB65613
 ID AAB65613 standard; Protein; 431 AA.
 AC AAB65613;

XX 27-MAR-2001 (first entry)

XX Novel protein kinase, SEQ ID NO: 139.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW immunological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

XX Homo sapiens.

XX WO200073469-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

XX 28-MAY-1999; 99US-0136503.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;

XX WPI; 2001-032161/04.

XX N-PSDB; AAF44639.

XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers.

XX Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC of kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.

XX Sequence 431 AA;

Query Match 100.0%; Score 103; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;

Matches	19;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	DPEETEEPVPNSIGKSPDS	19						
DB	386	DPEETEEPVPNSIGKSPDS	404						
RESULT 11									
AA65614									
ID	AA65614	standard; Protein;	430 AA.						
XX	AA65614;								
XX	27-MAR-2001	(first entry)							
XX	Novel protein kinase, SEQ ID NO: 140.								
XX	Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;								
XX	immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;								
XX	dermatological; antidiabetic; antifertility; gene therapy; vaccine;								
XX	immune disorder; cardiovascular disease; neurodegenerative disease;								
XX	cancer; autoimmune disorder; stroke; inflammatory bowel disease;								
XX	inflammatory pelvic disease; multiple sclerosis; psoriasis.								
XX	Mus musculus.								
XX	WO200073469-A2.								
XX	07-DEC-2000.								
XX	26-MAY-2000; 2000WO-US14842.								
XX	28-MAY-1999; 99US-0136503.								
XX	(SUGC-) SUGEN INC.								
XX	Plowman GD, Martinez R, Whyte D, Sudersanam S;								
XX	WPI; 2001-032161/04.								
XX	N-PSDR: AAF44640.								
PT	Nucleic acids encoding kinase polypeptides, useful for diagnosing and								
PT	treating immune-related diseases and disorders, cardiovascular disease,								
PT	neurodegenerative diseases and/or cancers.								
XX	Claim 10; Fig 1; 310pp; English.								
CC	The present sequence is a novel protein kinase. The novel protein kinases								
CC	and the nucleic acids that encode them may be used in the treatment and								
CC	diagnosis of diseases associated with inappropriate kinase expression								
CC	such as immune-related diseases and disorders, cardiovascular disease,								
CC	neurodegenerative diseases and/or cancers. The nucleic acids and								
CC	complementary sequences may also be used as DNA probes in diagnostic								
CC	assays. The kinase polypeptides may be used as antigens in the production								
CC	of antibodies of kinase expression and activity. Anti-kinase antibodies								
CC	and kinase antagonists may also be used to down regulate kinase								
CC	expression and activity. Diseases related to kinase expression and								
CC	activity include rheumatoid arthritis, atherosclerosis, autoimmune								
CC	disorders, complications of organ transplantation, myocardial infarction,								
CC	immune disorders, cardiomyopathies, strokes, renal failure,								
CC	oxidative stress related disorders, chronic inflammatory bowel disease,								
CC	chronic inflammatory pelvic disease, multiple sclerosis, asthma,								
CC	osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and								
CC	reproductive disorders.								
XX	Sequence 430 AA;								
Query Match	92.24; Score 95; DB 22; Length 430;								
Best Local Similarity	89.5%; Pred. No. 2.3e-06;								
Matches	17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;								
QY	1	DPEETEEPVPNSIGKSPDS	19						

Db	385	DPEETEEPVPSSIGKSPDS	403						
RESULT 12									
AA624116									
ID	AA624116	standard; Protein;	431 AA.						
XX	AA624116;								
XX	29-JAN-2001	(first entry)							
XX	Rat serum and glucocorticoid regulated kinase protein (RSGK).								
XX	Human; STK3; serine threonine kinase; HSGK; RSGK;								
XX	serum and glucocorticoid regulated kinase.								
XX	Rattus sp.								
XX	CNI259573-A.								
XX	12-JUL-2000.								
XX	29-OCT-1998; 98CN-0123822.								
XX	29-OCT-1998; 98CN-0123822.								
XX	(UYFU-) UNIV FUDAN.								
XX	Yu L, Fu Q, Zhao Y;								
XX	WPI; 2000-587991/56.								
XX	New human serine threonine protein kinase, its code sequence,								
XX	preparation, and use.								
XX	Example 3; Fig 4; 29pp; Chinese.								
XX	The present invention describes human serine threonine kinase STK3,								
XX	which is a new member of the serine/threonine protein kinase family.								
XX	Human STK3 shares homology with human serum and glucocorticoid								
XX	regulated kinase (HSGK) and rat SGK (RSGK). The present sequence								
XX	represents the rat RSGK protein sequence from the present invention.								
XX	Sequence 431 AA;								
Query Match	92.24; Score 95; DB 21; Length 431;								
Best Local Similarity	89.5%; Pred. No. 2.3e-06;								
Matches	17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;								
QY	1	DPEETEEPVPNSIGKSPDS	19						
DB	386	DPEETEEPVPSSIGKSPDS	404						
RESULT 13									
AA93530									
ID	AA93530	standard; Protein;	431 AA.						
XX	AA93530;								
XX	25-SEP-2000	(first entry)							
XX	A rat serum and glucocorticoid induced protein kinase.								
XX	Protein kinase; PKh1; PKh2; Ypk1; Yrk2; protein kinase B-alpha;								
XX	serum and glucocorticoid induced protein kinase; SGK; PKBalpha;								
XX	3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;								
XX	thrush; cancer; diabetes; obesity; antifungal; Candida infection.								
XX	Rattus sp.								
XX	WO2000036135-A2.								

PD 22-JUN-2000.
 XX
 XX 14-DEC-1999; 99WO-GB04228.
 XX
 PR 14-DEC-1998; 98US-0112114.
 XX
 XX (MEDI-) MEDICAL RES COUNCIL.
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Thorner JW, Alessi DR, Tortance PD, Casamayor A;
 PI WPI: 2000-442391/38.
 XX N-PSDB: AAA46592.
 DR
 XX
 XX Screening method identifying compounds which modulate protein kinase
 PT activity for use in treating fungal infections and cancer -
 PT
 PS Disclosure; Fig 13; 155pp; English.
 XX
 XX The present sequence represents a rat serum and glucocorticoid induced
 CC protein kinase (SGK). The specification describes a screening method
 CC to identify a compound which modulates the activity of protein kinases
 CC from different sources, using host yeast cells. The method is used to
 CC identify a compound which modulates (inhibits) the activity of a
 CC protein kinase. Pkhl or Pkh2 phosphorylate and activate Ypk1, Yrk2,
 CC SGK or protein kinase B-alpha (PKBalpha). 3-phosphoinositide-dependent
 CC protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
 CC Yrk2 or SGK but not PKBalpha or p70S6 kinase. Compounds identified by
 CC the methods are used to treat fungal infections e.g. thrush, and to
 CC treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the
 CC activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be
 CC used in the treatment of diabetes or obesity, and compounds which
 CC inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK
 CC (Ypk1 or Yrk2) can be used as an antifungal agent to treat Candida
 CC infections, e.g. thrush.
 XX
 XX Sequence 431 AA;
 SQ

Query Match 92.2%; Score 95; DB 21; Length 431;
 Best Local Similarity 89.5%; Pred. No. 2.3e-06;
 Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 DPETEEVPVNSIGKSPDS 19
 DB 386 DPETEEVPVNSIGKSPDS 404
 |||||

RESULT 14
 AAB65615
 ID AAB65615 standard; Protein: 244 AA.
 XX
 AC AAB65615;
 XX
 XX 27-MAR-2001 (first entry)
 DT
 XX Novel protein kinase, SEQ ID NO: 141.
 DE
 XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 XX Mus musculus.
 OS
 XX
 XX WO200073469-A2.
 PN
 XX
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US14842.
 PF
 XX 28-MAY-1999; 99US-0136503.
 PR

XX (SUGE-) SUGEN INC.
 XX
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 DR WPI: 2001-032161/04.
 DR N-PSDB: AAF44641.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 XX Claim 10; Fig 1; 310pp; English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 244 AA;
 Query Match 62.1%; Score 64; DB 22; Length 244;
 Best Local Similarity 63.2%; Pred. No. 0.078;
 Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DPETEEVPVNSIGKSPDS 19
 DB 200 DPETEEVPVNSIGKSPDS 218
 |||||

RESULT 15
 AAB99838
 ID AAB99838 standard; Protein: 319 AA.
 XX
 AC AAB99838;
 XX
 XX 20-SEP-2001 (first entry)
 DT
 XX AGC protein kinase family member SGK2 protein sequence.
 DE
 XX Protein kinase: identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX
 XX Homo sapiens.
 OS Synthetic.
 XX
 XX WO200144497-A2.
 PN
 XX 21-JUN-2001.
 PD
 XX 04-DEC-2000; 2000WO-GB04598.
 PF
 XX 02-DEC-1999; 99US-0168559.
 PR
 XX (UTDU-) UNIV DUNDEE.
 PA
 XX Alessi D, Biondi R;
 PI
 XX WPI: 2001-390252/41.
 DR

XX
PT
PT
PT
PT
XX
XX
PS
XX
XX

Identifying modulators of protein kinase (PK) activity, useful in developing drugs for treating cancer or diabetes, by measuring the ability of the compound to modulate or mimic the interaction of PK with interacting polypeptides -

Disclosure: Fig 16; 180pp; English.

The present invention describes a method for identifying a compound that modulates protein kinase activity. The method comprises measuring the ability of the compound to inhibit, promote or mimic the interaction of a hydrophobic pocket-containing protein kinase with an interacting polypeptide. The interacting polypeptide interacts with the hydrophobic pocket of the protein kinase and/or comprises the amino acid sequence Phe/Tyr-Xaa-Xaa-Phe/Tyr (1). The method is useful in screening assays for developing pharmaceutical compounds or drugs. Compounds, polypeptides or polynucleotides from the present invention are useful in medicine, particularly in the manufacture of a medicament for treating a patient in need of modulation of signalling by a hydrophobic pocket-containing protein kinase. Specifically, the patient has cancer or diabetes or is in need of inhibition of apoptosis, e.g. a patient suffering from tissue injury or ischaemic injury, including stroke. The compound or composition is also useful for inhibiting the degree or rate of phosphorylation by the protein kinase. The interacting polypeptide or compound is useful in methods of stabilising a hydrophobic pocket-containing protein kinase, where the protein kinase is exposed to the compound or polypeptide. AAB99786 to AAB99847 represent amino acid sequences, and AAB44210 and AAB44211 represent oligonucleotide sequences, used in the exemplification of the present invention.

XX
SQ

Sequence 319 AA;

Query Match 62.1%; Score 64; DB 22; Length 319;

Best Local Similarity 63.2%; Pred. No. 0.11;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DPFTTEPVPNSIGKSPDS 19

Db 276 DPFTQGANVSKSIGCTPDT 294

Search completed: August 11, 2003, 08:18:52
Job time : 7.03433 secs

Tue Aug 12 10:56:05 2003

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:08:14 ; Search time 5.08154 Seconds
(without alignments)
499.774 Million cell updates/sec

Title: US-10-000-039A-4

Perfect score: 85

Sequence: 1 EAFGRSYAPPTDSFL 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq19Jun03:*

1:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSL1/qcgcdata/geneseq/geneseq-emb1/AA2003.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	16	19 AAW77219	Human cell-volume
2	85	100.0	321	22 AAB95836	AGC protein kinase
3	85	100.0	373	22 AAB95815	AGC protein kinase
4	85	100.0	388	20 AAY48573	Human breast tumou
5	85	100.0	431	19 AAW77217	Human cell-volume
6	85	100.0	431	19 AAW54025	Human protein kina
7	85	100.0	431	20 AAW90139	Human sgk protein.
8	85	100.0	431	21 AAB24115	Human serum and gl
9	85	100.0	431	21 AAY95279	Human serum and gl

10	85	100.0	431	22 AAB65613	Novel protein kina
11	80	94.1	430	22 AAB65614	Novel protein kina
12	79	92.9	431	21 AAB24116	Rat serum and gl
13	79	92.9	431	21 AAY93530	A rat serum and gl
14	58.5	68.8	276	22 AAB97658	Novel central nerv
15	58.5	68.8	276	22 AAO17104	Novel signal trans
16	58.5	68.8	322	22 AAB95837	AGC protein kinase
17	58.5	68.8	429	21 AAB24114	Human serum and gl
18	58.5	68.8	429	21 AAY95278	Human serum and gl
19	58.5	68.8	446	22 AAB65616	Novel protein kina
20	58.5	68.8	464	22 AAW93732	Human polypeptide,
21	58.5	68.8	496	22 AAW87772	Human serum glucoc
22	58.5	68.8	496	22 AAO87228	Novel central nerv
23	58.5	68.8	496	22 AAO87382	Novel central nerv
24	58.5	68.8	496	22 AAM38834	Human polypeptide
25	58.5	68.8	496	22 AAM40620	Human polypeptide
26	56	65.9	244	22 AAB65615	Novel protein kina
27	56	65.9	319	22 AAB95838	AGC protein kinase
28	56	65.9	367	21 AAY95275	Human serum and gl
29	56	65.9	367	22 AAO28087	Novel human secret
30	56	65.9	367	22 AAB65708	Novel protein kina
31	56	65.9	367	23 AAE22765	Human serum and gl
32	56	65.9	382	22 AAM25594	Human protein sequ
33	56	65.9	427	21 AAY95276	Human serum and gl
34	51	60.0	9	21 AAY93662	SGK sequence for p
35	48	56.5	367	21 AAY95277	Mouse serum and gl
36	46	54.1	112	22 ABB43856	Human liver peptid
37	46	54.1	112	22 ABB29840	Peptide #2491 enco
38	46	54.1	112	22 ABB35021	Peptide #2527 enco
39	46	54.1	112	22 ABB20432	Protein #2431 enco
40	46	54.1	112	22 AAM55836	Human brain expres
41	46	54.1	112	22 AAM68209	Human bone marrow
42	46	54.1	112	22 AAM16027	Peptide #2461 enco
43	46	54.1	112	22 AAM28529	Peptide #2566 enco
44	46	54.1	112	22 AAM03758	Peptide #2440 enco
45	46	54.1	112	23 ABB33737	Human peptide enco

ALIGNMENTS

RESULT 1
ID AAW77219 standard; Peptide: 16 AA.
XX
AC AAW77219;
XX
DT 16-NOV-1998 (first entry)
XX
DE Human cell-volume regulating kinase h-sgk peptide fragment 2.
XX
KW Human; cell-volume; kinase; h-sgk; diabetes mellitus;
XX renal insufficiency; inflammation; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN EP861896-A2.
XX
PD 02-SEP-1998.
XX
PF 27-JAN-1998; 98EP-0101338.
XX
PR 28-FEB-1997; 97DE-1008173.
XX
(DADE-) DADE BEHRING MARBURG GMBH.
XX
Lang F, Waldegger S;
XX
WPI: 1998-449109/39.
XX
New nucleic acid encoding cell-volume regulating kinase h-sgk and
XX related proteins - used for diagnosis and treatment of diseases
XX involving changes in cell volume, e.g. renal insufficiency.

PT inflammation, infections etc.
 PS Claim 7; Page 9; 15pp; German.
 CC The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natremia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.
 XX
 XX SQ Sequence 16 AA:
 Query Match 100.0%; Score 85; DB 19; Length 16;
 Best Local Similarity 100.0%; Pred. No. 8.6e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 1 EAFLGFSYAPPTDSFL 16
 |||||
 RESULT 2
 AAB99836
 ID AAB99836 standard; Protein: 321 AA.
 XX AC AAB99836;
 XX DT 20-SEP-2001 (first entry)
 XX DE AGC protein kinase family member SGK1 protein sequence.
 XX KW Protein kinase; identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN WO200144497-A2.
 XX PD 21-JUN-2001.
 XX PF 04-DEC-2000; 2000WO-GB04598.
 XX PR 02-DEC-1999; 99US-0168559.
 XX PA (UYDU-) UNIV DUNDEE.
 XX PI Alessi D, Biondi R;
 XX WPI; 2001-390252/41.
 XX PT Identifying modulators of protein kinase (PK) activity, useful in
 PT developing drugs for treating cancer or diabetes, by measuring the
 PT ability of the compound to modulate or mimic the interaction of PK with
 PT interacting polypeptides -
 XX Disclosure; Fig 16; 180pp; English.
 XX The present invention describes a method for identifying a compound that
 CC modulates protein kinase activity. The method comprises measuring the
 CC ability of the compound to inhibit, promote or mimic the interaction of
 CC a hydrophobic pocket-containing protein kinase with an interacting
 CC polypeptide. The interacting polypeptide interacts with the hydrophobic
 CC pocket of the protein kinase and/or comprises the amino acid sequence
 CC Phe/Tyr-Xaa-Xaa-Phe/Tyr (1). The method is useful in screening assays
 CC for developing pharmaceutical compounds or drugs. Compounds, polypeptides
 CC or polynucleotides from the present invention are useful in medicine,
 CC particularly in the manufacture of a medicament for treating a patient
 CC in need of modulation of signalling by a hydrophobic pocket-containing

CC particularly in the manufacture of a medicament for treating a patient
 CC in need of modulation of signalling by a hydrophobic pocket-containing
 CC protein kinase. Specifically, the patient has cancer or diabetes or is
 CC in need of inhibition of apoptosis, e.g. a patient suffering from tissue
 CC injury or ischaemic injury, including stroke. The compound or
 CC composition is also useful for inhibiting the degree or rate of
 CC phosphorylation by the protein kinase. The interacting polypeptide or
 CC compound is useful in methods of stabilising a hydrophobic pocket-
 CC containing protein kinase, where the protein kinase is exposed to the
 CC compound or polypeptide. AAB99786 to AAB99847 represent amino acid
 CC sequences, and AAB44210 and AAB44211 represent oligonucleotide sequences,
 CC used in the exemplification of the present invention.
 XX
 XX SQ Sequence 321 AA:
 Query Match 100.0%; Score 85; DB 22; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 306 EAFLGFSYAPPTDSFL 321
 |||||
 RESULT 3
 AAB99815
 ID AAB99815 standard; Protein: 373 AA.
 XX AC AAB99815;
 XX DT 20-SEP-2001 (first entry)
 XX DE AGC protein kinase family member SGK protein sequence.
 XX KW Protein kinase; identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX OS Homo sapiens.
 XX OS Synthetic.
 XX PN WO200144497-A2.
 XX PD 21-JUN-2001.
 XX PF 04-DEC-2000; 2000WO-GB04598.
 XX PR 02-DEC-1999; 99US-0168559.
 XX PA (UYDU-) UNIV DUNDEE.
 XX PI Alessi D, Biondi R;
 XX WPI; 2001-390252/41.
 XX PT Identifying modulators of protein kinase (PK) activity, useful in
 PT developing drugs for treating cancer or diabetes, by measuring the
 PT ability of the compound to modulate or mimic the interaction of PK with
 PT interacting polypeptides -
 XX Disclosure; Fig 15; 180pp; English.
 XX The present invention describes a method for identifying a compound that
 CC modulates protein kinase activity. The method comprises measuring the
 CC ability of the compound to inhibit, promote or mimic the interaction of
 CC a hydrophobic pocket-containing protein kinase with an interacting
 CC polypeptide. The interacting polypeptide interacts with the hydrophobic
 CC pocket of the protein kinase and/or comprises the amino acid sequence
 CC Phe/Tyr-Xaa-Xaa-Phe/Tyr (1). The method is useful in screening assays
 CC for developing pharmaceutical compounds or drugs. Compounds, polypeptides
 CC or polynucleotides from the present invention are useful in medicine,
 CC particularly in the manufacture of a medicament for treating a patient
 CC in need of modulation of signalling by a hydrophobic pocket-containing

CC protein kinase. Specifically, the patient has cancer or diabetes or is
 CC in need of inhibition of apoptosis, e.g. a patient suffering from tissue
 CC injury or ischemic injury, including stroke. The compound or
 CC composition is also useful for inhibiting the degree or rate of
 CC phosphorylation by the protein kinase. The interacting polypeptide or
 CC compound is useful in methods of stabilising a hydrophobic pocket-
 CC containing protein kinase, where the protein kinase is exposed to the
 CC compound or polypeptide. AAB99786 to AAB99847 represent amino acid
 CC sequences, and AAH44210 and AAH44211 represent oligonucleotide sequences,
 CC used in the exemplification of the present invention.

XX SQ Sequence 373 AA;

Query Match 100.0%; Score 85; DB 22; Length 373;
 Best Local Similarity 100.0%; Pred. No. 2e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
 DB 358 EAFLGFSYAPPTDSFL 373

RESULT 4

AA48573
 ID AAY48573 standard; Protein; 388 AA.

XX AC AAY48573;

XX DT 08-DEC-1999 (first entry)

XX DE Human breast tumour-associated protein 34.

XX KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;
 XX KW treatment; tumour; cytostatic; medicament.

XX OS Homo sapiens.

XX PN DE19813839-A1.

XX PO 23-SEP-1999.

XX PF 20-MAR-1998; 98DE-1013839.

XX PR 20-MAR-1998; 98DE-1013839.

XX PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;

XX DR WPI: 1999-528981/45.

XX DR N-PSDB; AA433647.

XX PT Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -

XX PS Claim 22; 157-158; 188pp; German.

XX CC This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AA48540-Y48617 represent protein
 CC fragments encoded by the expressed sequence tags described in the method
 CC of the invention.

XX SQ Sequence 388 AA;

Query Match 100.0%; Score 85; DB 20; Length 388;

Best Local Similarity 100.0%; Pred. No. 2e-05;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
 DB 373 EAFLGFSYAPPTDSFL 388

RESULT 5

AAW77217
 ID AAW77217 standard; Protein; 431 AA.

XX AC AAW77217;

XX DT 16-NOV-1998 (first entry)

XX DE Human cell-volume regulating kinase h-sgk.

XX KW Human; cell-volume; kinase; h-sgk; diabetes mellitus;
 XX KW renal insufficiency; inflammation; Alzheimer's disease.

XX OS Homo sapiens.

XX PN EP861896-A2.

XX PD 02-SEP-1998.

XX PF 27-JAN-1998; 98EP-0101338.

XX PR 28-FEB-1997; 97DE-1008173.

XX PA (DADE-) DADE BEHRING MARBURG GMBH.

XX PI Lang F, Waldegger S;

XX DR WPI: 1998-449109/39.

XX DR N-PSDB; AAV48311.

XX PT New nucleic acid encoding cell-volume regulating kinase h-sgk and
 PT related proteins - used for diagnosis and treatment of diseases
 PT involving changes in cell volume, e.g. renal insufficiency,
 PT inflammation, infections etc.

XX PS Disclosure; Fig 1; 15pp; German.

XX CC The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 85; DB 19; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.3e-05;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16

DB 416 EAFLGFSYAPPTDSFL 431

RESULT 6

AAW54025
 ID AAW54025 standard; Protein; 431 AA.

XX AC AAW54025;

XX DT 31-JUL-1998 (first entry)

XX DE Human protein kinase HPK-3.

XX Protein kinase; human; HPK; signalling cascade; kinase expression;
 KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;
 KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;
 KW lymphoma; therapy.
 XX Homo sapiens.
 XX WO9611234-A2.
 PN WO9611234-A2.
 XX 19-MAR-1998.
 PD 10-SEP-1997; 97WO-USI5923.
 PF 12-SEP-1996; 96US-0712709.
 PR (INCY-) INCYTE PHARM INC.
 XX Au-Young J, Guegler KJ, Hawkins PR;
 PI WPI; 1998-207394/18.
 DR N-PSDB; AAV23833.
 UR New isolated human protein kinase(s) - used to develop products for
 PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
 PT asthma, multiple sclerosis or rheumatoid arthritis
 XX Claim 23; Fig 3; 75pp; English.
 PS This sequence is a human protein kinase (HPK) of the invention. The
 CC HPK protein can be used to develop products for studying signalling
 CC cascades in various cells and tissues, diagnosing disease and selecting
 CC inhibitors or drugs with the potential to intervene in various disorders
 CC or diseases in which altered kinase expression is implicated. The
 CC products can be used to e.g. reverse memory loss such as due to
 CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
 CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
 CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
 CC lymphocytic leukaemias or lymphomas.
 XX Sequence 431 AA:
 SQ
 Query Match 100.0%; Score 85; DB 19; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 Db 416 EAFLGFSYAPPTDSFL 431
 RESULT 7
 AAW90139
 ID AAW90139 standard; Protein: 431 AA.
 AC AAW90139;
 XX 15-MAR-1999 (first entry)
 DT Human sgk protein.
 DE Serine glucocorticoid regulated kinase; sgk; human; treatment: inhibitor;
 KW serine/threonine protein kinase family; antagonist: diabetic nephropathy;
 KW chronic renal failure; inflammation; Alzheimers disease; wound.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 XX Key Misc-difference 381
 FT /note= "encoded by GAG"
 XX EP887081-A2.
 PN

PD 30-DEC-1998.
 XX 27-MAY-1998; 98EP-0304189.
 PF 27-JUN-1997; 97US-0051124.
 PR (SMIK) SMITHKLINE BEECHAM CORP.
 XX Kumar JM;
 XX NPI; 1999-047627/05.
 DR N-PSDB; AAV74190.
 UR Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 PT disease - by administration of nucleic acids and antagonists which
 PT inhibit activity or expression of human serum glucocorticoid
 PT regulated kinase (sgk), a serine/threonine protein kinase
 XX Disclosure; page 16-17; 17pp; English.
 PS This sequence represents a novel human serum glucocorticoid regulated
 CC kinase (sgk) protein which is a member of the serine/threonine protein
 CC kinase family. This protein is used for the treatment of a subject having
 CC need to inhibit/antagonise activity or expression of human sgk
 CC polypeptide e.g. for the treatment of chronic renal failure, diabetic
 CC nephropathy, inflammation, Alzheimers disease and wounds.
 XX Sequence 431 AA:
 SQ
 Query Match 100.0%; Score 85; DB 20; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 Db 416 EAFLGFSYAPPTDSFL 431
 RESULT 8
 AAB24115
 ID AAB24115 standard; Protein: 431 AA.
 AC AAB24115;
 XX 29-JAN-2001 (first entry)
 DT Human serum and glucocorticoid regulated kinase protein (HSRK).
 DE Human; STK3; serine threonine kinase; HSKG; RSKG;
 KW serum and glucocorticoid regulated kinase.
 XX Homo sapiens.
 OS CNL259573-A.
 PN CNL259573-A.
 XX 12-JUL-2000.
 PD 29-OCT-1998; 98CN-0123822.
 PF 29-OCT-1998; 98CN-0123822.
 PR (UYFU-) UNIV FUDAN.
 XX Yu L, Fu Q, Zhao Y;
 PI WPI; 2000-587991/56.
 DR New human serine threonine protein kinase, its code sequence,
 XX preparation, and use
 PT Example 3; Fig 3; 29pp; Chinese.
 PS The present invention describes human serine threonine kinase STK3,
 CC

CC which is a new member of the serine/threonine protein kinase family.
 CC Human SGK3 shares homology with human serum and glucocorticoid
 CC regulated kinase (HSGK) and rat SGK (RSGK). The present sequence
 CC represents the human HSGK protein sequence from the present invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 85; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 DB 416 EAFLGFSYAPPTDSFL 431
 RESULT 9
 ID AAY95279 standard; Protein: 431 AA.
 XX
 AC AAY95279;
 DT 12-SEP-2000 (first entry)
 XX
 DE Human serum and glucocorticoid-induced protein kinase.
 XX
 KW Serum and glucocorticoid-induced protein kinase; SGK;
 KW human; phosphorylation; cancer; diabetes; ischaemia; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 256
 FT /note= "O-phosphorylated"
 FT Modified-site 422
 FT /note= "O-phosphorylated"
 XX
 WO200035946-A1.
 XX
 22-JUN-2000.
 XX
 PF 14-DEC-1999; 99WO-CB04232.
 XX
 PR 14-DEC-1998; 98US-0112217.
 XX
 PR 19-AUG-1999; 99GB-0019676.
 XX
 PA (UYDU-) UNIV DUNDSE.
 XX
 PI Cohen P, Kobayashi T, Deak M;
 XX
 WPI; 2000-442364/38.
 XX
 PT Modulation of serum and glucocorticoid-induced protein kinase (SGK) by
 PT phosphorylation with 3-phosphoinositide-dependent protein kinase-1
 PT (PDK1) or dephosphorylation, useful for treatment of cancer, diabetes
 PT and ischemic diseases -
 XX
 PS Disclosure; Fig 13; 127pp; English.
 XX
 CC The present sequence is that of human serum and
 CC glucocorticoid-induced protein kinase (SGK). SGK was initially
 CC identified as a glucocorticoid and osmotic stress-responsive gene.
 CC It is activated by phosphorylation of its Thr-308 and Ser-472
 CC residues. Novel isoforms of SGK, i.e. SGK2 and SGK3 (see
 CC AAY95275-78), have been identified. The invention provides methods
 CC of activating SGK activity by phosphorylation using
 CC 3-phosphoinositide-dependent protein kinase-1 (PDK1), and of
 CC reducing the activity of SGK by dephosphorylation. The invention
 CC also provides a method of identifying a compound that modulates the
 CC activity of SGK. Such compounds are useful for treating patients
 CC requiring modulation of SGK, such as patients with cancer, diabetes
 CC or ischaemic disease.

SQ Sequence 431 AA;
 Query Match 100.0%; Score 85; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDSFL 16
 DB 416 EAFLGFSYAPPTDSFL 431
 RESULT 10
 ID AAB65613 standard; Protein: 431 AA.
 XX
 AC AAB65613;
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 139.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14842.
 XX
 PR 28-MAY-1999; 99US-0136503.
 XX
 PA (SUGF-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 WPI; 2001-032161/04.
 DR N-PSOB; AAF44639.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 PS Claim 10; Fig 1; 310pp; English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma, and
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 85; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-05;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFGLGSYAPPTDSFL 16
 DB 416 EAFGLGSYAPPTDSFL 431

RESULT 11
 AAB65614
 ID AAB65614 standard; Protein; 430 AA.
 XX
 AC AAB65614;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 140.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Mus musculus.
 XX
 PN W0200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14842.
 XX
 PR 28-MAY-1999; 99US-0136503.
 XX
 PA (SUGF-) SUGEN INC.
 XX
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 DR WPI; 2001-032161/04.
 DR N-PSDB; AAE44640.
 XX
 PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -

Claim 10; Fig 1; 310pp; English.

The present sequence is a novel protein kinase. The novel protein kinases and the nucleic acids that encode them may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders.

Sequence 430 AA:
 Query Match 94.1%; Score 80; DB 22; Length 430;
 Best Local Similarity 93.8%; Pred. No. 0.00014;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFGLGSYAPPTDSFL 16
 DB 416 EAFGLGSYAPPTDSFL 431

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFGLGSYAPPTDSFL 16
 DB 416 EAFGLGSYAPPTDSFL 431

RESULT 12
 AAB24116
 ID AAB24116 standard; Protein; 431 AA.
 XX
 AC AAB24116;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Rat serum and glucocorticoid regulated kinase protein (RSGK).
 XX
 KW Human; STK3; serine threonine kinase; HSGK; RSGK;
 KW serum and glucocorticoid regulated kinase.
 XX
 OS Rattus sp.
 XX
 PN CN1259573-A.
 XX
 PD 12-JUL-2000.
 XX
 PF 29-OCT-1998; 98CN-0123822.
 XX
 PR 29-OCT-1998; 98CN-0123822.
 XX
 PA (UYFU-) UNIV FUDAN.
 XX
 PI Yu L, Fu Q, Zhao Y;
 XX
 DR WPI; 2000-587991/56.
 XX
 PT New human serine threonine protein kinase, its code sequence,
 PT preparation, and use -

Example 3; Fig 4; 29pp; Chinese.

The present invention describes human serine threonine kinase STK3, which is a new member of the serine/threonine protein kinase family. Human STK3 shares homology with human serum and glucocorticoid regulated kinase (HSGK) and rat SK (RSGK). The present sequence represents the rat RSGK protein sequence from the present invention.

Sequence 431 AA:
 Query Match 92.9%; Score 79; DB 21; Length 431;
 Best Local Similarity 93.8%; Pred. No. 0.0002;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFGLGSYAPPTDSFL 16
 DB 416 EAFGLGSYAPPTDSFL 431

RESULT 13
 AAY93530
 ID AAY93530 standard; Protein; 431 AA.
 XX
 AC AAY93530;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE A rat serum and glucocorticoid induced protein kinase.
 XX
 KW Protein kinase; PKh1; PKh2; Ypk1; Yrk2; protein kinase B-alpha;
 KW serum and glucocorticoid induced protein kinase; SK; PKAalpha;
 KW 3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
 KW thrush; cancer; diabetes; obesity; antifungal; Candida infection.
 XX
 OS Rattus sp.
 XX
 PN W0200036135-A2.
 XX

```

PD 22-JUN-2000.
XX
XX 14-DEC-1999; 99WO-GB04228.
XX
XX 14-DEC-1998; 98US-0112114.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Thorne JW, Alessi DR, Torrance PD, Casamayor A;
XX
XX WPI: 2000-442381/38.
XX N-PSDB: AAR46592.
XX
XX Screening method identifying compounds which modulate protein kinase
XX activity for use in treating fungal infections and cancer -
XX
XX Disclosure; Fig 13: 155pp; English.
XX
XX The present sequence represents a rat serum and glucocorticoid induced
XX protein kinase (SGK). The specification describes a screening method
XX to identify a compound which modulates the activity of protein kinases
XX from different sources, using host yeast cells. The method is used to
XX identify a compound which modulates (inhibits) the activity of a
XX protein kinase. Pkhl or Pkh2 phosphorylate and activate Ypk1, Yrk2,
XX SGK or protein kinase B-alpha (PKBalpha). 3-phosphoinositide-dependent
XX protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
XX Yrk2 or SGK but not PKBalpha or p70S6 kinase. Compounds identified by
XX the methods are used to treat fungal infections e.g. thrush, and to
XX treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the
XX activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be
XX used in the treatment of diabetes or obesity, and compounds which
XX inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK
XX (Ypk1 or Yrk2) can be used as an antifungal agent to treat Candida
XX infections, e.g. thrush.
XX
XX Sequence 431 AA;
XX
XX Query Match 92.9%; Score 79; DB 21; Length 431;
XX Best Local Similarity 93.8%; Pred. NO. 0.0002;
XX Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EARLGFYSYAPPTDSEFL 16
XX | | | | | | | | | | | | | |
XX DB 416 EARLGFYSYAPPTDSEFL 431
XX
XX RESULT 14
XX AAU87658
XX ID AAU87658 standard; Protein; 276 AA.
XX
XX AC AAU87658;
XX
XX DT 05-JUN-2002 (first entry)
XX
XX DE
XX
XX KW Central nervous system protein #568.
XX
XX KW hyperproliferative disorder; CNS; autoimmune disease; rheumatoid arthritis;
XX cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
XX nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
XX acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
XX adenocarcinoma; reproductive system disorder; testicular feminisation;
XX endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
XX respiratory disorder; renal disorder; kidney failure; blood disorder;
XX myocardial infarction; wound healing; cell proliferation; skin aging;
XX food additive; food preservative; gene therapy.
XX
XX OS Homo sapiens.
XX
XX PN WO200155318-A2.
XX
XX PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01332.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX 04-FEB-2000; 2000US-0180828.
XX
XX 24-FEB-2000; 2000US-0184664.
XX
XX 02-MAR-2000; 2000US-0186350.
XX
XX 16-MAR-2000; 2000US-0189874.
XX
XX 17-MAR-2000; 2000US-0190076.
XX
XX 18-APR-2000; 2000US-0198123.
XX
XX 19-MAY-2000; 2000US-0205515.
XX
XX 07-JUN-2000; 2000US-0209467.
XX
XX 28-JUN-2000; 2000US-0214886.
XX
XX 30-JUN-2000; 2000US-0215135.
XX
XX 07-JUL-2000; 2000US-0216647.
XX
XX 07-JUL-2000; 2000US-0216880.
XX
XX 11-JUL-2000; 2000US-0217487.
XX
XX 14-JUL-2000; 2000US-0217496.
XX
XX 26-JUL-2000; 2000US-0218290.
XX
XX 26-JUL-2000; 2000US-0220963.
XX
XX 14-AUG-2000; 2000US-0220964.
XX
XX 14-AUG-2000; 2000US-0224518.
XX
XX 14-AUG-2000; 2000US-0224519.
XX
XX 14-AUG-2000; 2000US-0225213.
XX
XX 14-AUG-2000; 2000US-0225214.
XX
XX 14-AUG-2000; 2000US-0225266.
XX
XX 14-AUG-2000; 2000US-0225267.
XX
XX 14-AUG-2000; 2000US-0225268.
XX
XX 14-AUG-2000; 2000US-0225270.
XX
XX 14-AUG-2000; 2000US-0225447.
XX
XX 14-AUG-2000; 2000US-0225757.
XX
XX 14-AUG-2000; 2000US-0225758.
XX
XX 18-AUG-2000; 2000US-0225759.
XX
XX 18-AUG-2000; 2000US-0226279.
XX
XX 22-AUG-2000; 2000US-0226681.
XX
XX 22-AUG-2000; 2000US-0226868.
XX
XX 22-AUG-2000; 2000US-0227182.
XX
XX 23-AUG-2000; 2000US-0227009.
XX
XX 30-AUG-2000; 2000US-0228924.
XX
XX 01-SEP-2000; 2000US-0229287.
XX
XX 01-SEP-2000; 2000US-0229343.
XX
XX 01-SEP-2000; 2000US-0229344.
XX
XX 01-SEP-2000; 2000US-0229345.
XX
XX 05-SEP-2000; 2000US-0229509.
XX
XX 05-SEP-2000; 2000US-0229513.
XX
XX 06-SEP-2000; 2000US-0230437.
XX
XX 06-SEP-2000; 2000US-0230438.
XX
XX 08-SEP-2000; 2000US-0231242.
XX
XX 08-SEP-2000; 2000US-0231243.
XX
XX 08-SEP-2000; 2000US-0231244.
XX
XX 08-SEP-2000; 2000US-0231413.
XX
XX 08-SEP-2000; 2000US-0231414.
XX
XX 08-SEP-2000; 2000US-0232080.
XX
XX 12-SEP-2000; 2000US-0232081.
XX
XX 14-SEP-2000; 2000US-0232397.
XX
XX 14-SEP-2000; 2000US-0232398.
XX
XX 14-SEP-2000; 2000US-0232399.
XX
XX 14-SEP-2000; 2000US-0232400.
XX
XX 14-SEP-2000; 2000US-0233063.
XX
XX 14-SEP-2000; 2000US-0233064.
XX
XX 21-SEP-2000; 2000US-0233065.
XX
XX 21-SEP-2000; 2000US-0234223.
XX
XX 21-SEP-2000; 2000US-0234274.
XX
XX 25-SEP-2000; 2000US-0234997.
XX
XX 25-SEP-2000; 2000US-0234998.
XX
XX 26-SEP-2000; 2000US-0235484.
XX
XX 27-SEP-2000; 2000US-0235834.
XX
XX 27-SEP-2000; 2000US-0235836.
XX
XX 29-SEP-2000; 2000US-0236327.
XX
XX 29-SEP-2000; 2000US-0236367.
XX
XX 29-SEP-2000; 2000US-0236368.

```

PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249254.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251858.
 PR 08-DEC-2000; 2000US-0251859.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR {HUMAN} HUMAN GENOME SCI INC.
 PR Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-581633/65.
 DR N-PSDB; ABK43988.

XX New isolated nucleic acid encoding a protein for diagnosing,
 PT preventing, treating or ameliorating medical conditions and used as
 PT food additives or preservatives -
 XX Claim 9; SEQ ID NO 1176; 837pp; English.
 XX The invention describes an isolated nucleic acid molecule (I) encoding a
 CC novel central nervous system protein. (I) and polypeptides (II) encoded
 CC by (I), are used to treat a medical condition and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiotensin, nervous system disorders e.g. Alzheimer's disease and
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukaemia, disorders involving neovascularisation e.g. malignancies, e.g.
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 68.8%; Score 58.5; DB 22; Length 276;
 Best Local Similarity 70.6%; Pred. NO. 0.24;
 Matches 12; Conservative 3; Mismatches 1; Indels 1; Gaps 1;
 Qy 1 EAFVGFSTAPPDPT-DSPL 16
 Db 260 DAFVGFSTAPPSEDLFL 276
 RESULT 15
 ID AAU17104 standard; Protein; 276 AA.
 XX AAU17104;
 AC AAU17104;
 XX 07-NOV-2001 (first entry)
 DE Novel signal transduction pathway protein, Seq ID 669.
 XX Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;
 KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
 KW immune system disorder; rheumatoid arthritis; inflammatory condition;
 KW organ transplant rejection; infection; hepatitis C; blood disorder;
 KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
 KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
 KW reproductive system; gastrointestinal; liver disorder; AIDS;
 XX acquired immune deficiency syndrome.
 OS Homo sapiens.
 XX WO200154733-A1.
 PN 02-AUG-2001.
 PD 17-JAN-2001; 2001WO-US01312.
 PF 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184564.
 PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.
XX N-PSDB; AAS27021.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
XX prognosing disorders related to the proteins, including cancers, immune
XX disorders and neuronal disorders.
PS Claim 1; SEQ ID No 569; 880pp; English.
XX

The invention relates to novel isolated polypeptides (I), and polynucleotides (II). (I) and the antibody to (I) are useful for diagnosing, preventing and treating diseases including immune system disorders (e.g. congenital and acquired immunodeficiencies, autoimmune disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ transplant rejections and graft versus host disease, infectious diseases (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and other blood-related disorders (sickle cell anaemia), myeloproliferative disorders, primary haematopoietic disorders, hyperproliferative disorders (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal disorders (e.g. Down syndrome), ischaemic injury (e.g. stroke), renal abnormalities (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders, dermatological disorders, wound healing, epithelial cell proliferation, endocrine disorders (e.g. Addison's disease), reproductive system disorders, gastrointestinal disorder (inflammatory disorders), liver disorders (cirrhosis), as stimulators of B-cell responsiveness to pathogens, activators of T-cells, to induce higher affinity antibodies, and as a means to induce tumour proliferation in pathological e.g. acquired immune deficiency syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction pathway protein, amino acid sequences of the invention.

Query Match 68.8%; Score 58.5; DB 22; Length 276;
Best Local Similarity 70.6%; Pred. No. 0.24;
Matches 12; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

Search completed: August 11, 2003, 08:18:52
Job time : 5.08154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein - protein search, using sw model
Run on: August 11, 2003, 08:05:45 : Search time 2.12876 Seconds
(without alignments)
722.816 Million cell updates/sec

Title: US-10-000-039A-4
Perfect score: 85
Sequence: 1 EAFGLGFSYAPPTDSFL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR76:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	92.9	431	2 A48094	serum and glucocorticoid-regulated kinase - rat
2	54	63.5	374	2 E90292	transmembrane protein (permease) [Imported] - Sulfolobus solfataricus
3	54	63.5	374	2 F90361	transmembrane protein (permease) [Imported] - Sulfolobus solfataricus
4	45	52.9	330	2 A85585	unknown protein en
5	45	52.9	330	2 G90734	hypothetical prote
6	43	50.6	162	2 T38904	conserved hypothet
7	43	50.6	283	2 S73233	probable lipoprote
8	43	50.6	444	2 D67557	major facilitator
9	43	50.6	481	2 J60377	p70 S6 kinase (EC
10	43	50.6	802	2 T45642	Fish metalloprotei
11	42.5	50.0	680	2 C82437	transketolase 1 VC
12	42	49.4	348	2 T26528	hypothetical prote
13	42	49.4	433	2 S74753	hypothetical prote
14	41.5	48.8	946	2 S54354	inter-alpha-inhibi
15	41	48.2	171	2 H71289	hypothetical prote
16	41	48.2	335	2 A75363	inositol monophosp
17	41	48.2	349	2 T32633	hypothetical prote
18	41	48.2	401	2 A80696	probable integral
19	41	48.2	403	2 F64923	probable membrane
20	41	48.2	403	2 A90925	probable transport
21	41	48.2	403	2 E85773	probable transport
22	41	48.2	412	2 A80080	tRNA adenylyltrans
23	41	48.2	567	2 T33400	protein Kinase C h
24	41	48.2	597	2 T33399	protein Kinase C h
25	41	48.2	704	1 S60117	transcription regu
26	40	47.1	202	2 A87636	hypothetical prote
27	40	47.1	272	2 E64248	hypothetical prote
28	40	47.1	293	2 AG2929	hypothetical prote
29	40	47.1	293	2 G98352	dipeptide ABC tran

ALIGNMENTS

RESULT 1

A48094
serum and glucocorticoid-regulated kinase - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 19-Dec-1997
C:Accession: A48094
R:Webster, M.K.; Goya, L.; Ge, Y.; Maiyar, A.C.; Firestone, G.L.
Mol. Cell. Biol. 13, 2031-2040, 1993
A:Title: Characterization of sgk, a novel member of the serine/threonine protein kinase family: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
A:Reference number: A48094; MUID:93204949; PMID:8455596
A:Accession: A48094
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-431 <WEB>
A:Experimental source: Con8.hd6 mammary epithelial tumor cells
A:Note: sequence extracted from NCBI backbone (NCBI:127618, NCBI:127619)
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP
F:96-355/Domain: protein kinase homology <KIN>
F:104-112/Region: protein kinase ATP-binding motif

Query Match 92.9%; Score 79; DB 2; Length 431;
Best Local Similarity 93.8%; Pred. No. 4e-05; Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFGLGFSYAPPTDSFL 16
|||||
DB 416 EAFGLGFSYAPPTDSFL 431

RESULT 2

E90292
transport membrane protein (permease) [Imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: E90292
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Claret, R.A.; Ragan, M.A.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: E90292
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <KUR>
A:Cross-references: GB:AE006641; NID:ql3814572; PIDN:AAK41596.1; GSPDB:GN00155
C:Genetics:
A:Gene: S901361

Query Match 63.5%; Score 54; DB 2; Length 374;
Best Local Similarity 60.0%; Pred. No. 0.4;

Matches 12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;
 QY 1 EAFL----GFSYAPPTDSFL 16
 |||| ||| ||||| :
 DB 290 EAFLGIFSFYSYRPTDSLI 309

RESULT 3
 F90361
 transport membrane protein (permease) [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: F90361
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
 arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: F90361
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-374 <KUR>
 A:Cross-references: GB:AE006641; NID:g13815231; PIDN:AAK42149.1; CSPDB:GN00155
 C:Genetics:
 A:Gene: SSO1958

Query Match 63.5%; Score 54; DB 2; Length 374;
 Best Local Similarity 60.0%; Pred. No. 0.4;
 Matches 12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 EAFL----GFSYAPPTDSFL 16
 |||| ||| ||||| :
 DB 290 EAFLGIFSFYSYRPTDSLI 309

RESULT 4
 A85585
 unknown protein encoded by prophage CP-933K [imported] - Escherichia coli (strain O157:H7)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: A85585
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; M01D:21074935; PMID:11206351
 A:Accession: A85585
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-330 <STO>
 A:Cross-references: GB:AE005174; NID:g12513756; PIDN:AAG55141.1; GSPDB:GN00145; UNGCP:209
 A:Experimental source: strain O157:H7, Substrain EDL933
 C:Genetics:
 A:Gene: Z0986

Query Match 52.9%; Score 45; DB 2; Length 330;
 Best Local Similarity 53.3%; Pred. No. 1.0;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSF 15
 | : ||| ||| :
 DB 158 EPIVVSFVSAPPTDSY 172

RESULT 5
 G90734
 hypothetical protein ECs0847 [imported] - Escherichia coli (strain O157:H7, Substrain R)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: G90734
 R:Hayashi, T.; Nakino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and a
 A:Reference number: A99629; M01D:21156231; PMID:11258796
 A:Accession: G90734
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-330 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA034270.1; PID:g13360306; GSPDB:GN00154
 A:Experimental source: strain O157:H7, Substrain RMD 0509952
 C:Genetics:
 A:Gene: ECs0847

Query Match 52.9%; Score 45; DB 2; Length 330;
 Best Local Similarity 53.3%; Pred. No. 1.0;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSF 15
 | : ||| ||| :
 DB 158 EPIVVSFVSAPPTDSY 172

RESULT 6
 T38904
 conserved hypothetical protein SPAC56E4.02c - fission yeast (Schizosaccharomyces pombe
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T38904
 R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z21813
 A:Accession: T38904
 A:Status: Preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-162 <CON>
 A:Cross-references: EMBL:Z99261; PIDN:CAH16393.1; GSPDB:GN000066; SMID:SPAC56E4.02c
 A:Experimental source: strain 972h-; cosmid c56E4
 C:Genetics:
 A:Gene: SPDB:SPAC56E4.02c
 A:Map position: 1

Query Match 50.6%; Score 43; DB 2; Length 162;
 Best Local Similarity 53.8%; Pred. No. 1.0;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LGFSYAPPTDSFL 16
 ||| ||| : : :
 DB 59 LGFDYAPPEISYI 71

RESULT 7
 S73523
 probable lipoprotein E09_ort283a - Mycoplasma pneumoniae (strain ATCC 29342)
 N:Alternate names: MG439 homologE09_ort283a
 C:Species: Mycoplasma pneumoniae
 C:Variety: ATCC 29342
 C>Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 C:Accession: S73523
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkkl, E.; Li, B.C.; Hertzmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
 A:Reference number: S73327; M01D:97105885; PMID:8948633
 A:Accession: S73523
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-283 <HIM>
 A:Cross-references: EMBL:AE000020; GB:U00089; NID:g1673852; PIDN:AAB95845.1; PID:g167
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: hypothetical protein MG440

Query Match 50.6%; Score 43; DB 2; Length 283;
 Best Local Similarity 50.0%; Pred. No. 18;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDS 14
: | | | | |
DB 204 DGFTFTYPTQTS 217

RESULT 8

major facilitator family transporter CC2485 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87557
R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
N.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
C:Cross-references: GB:AE005673; NID:g13424038; PIDN:AAK24456.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2485

Query Match 50.6%; Score 43; DB 2; Length 444;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDS 15
: | | | | |
DB 241 KAFYGYAPPIASF 255

RESULT 9

p70 S6 kinase (EC 2.7.-.-) - human
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
C:Accession: JE0377
R:Salton, M.; ten Dijke, P.; Miyazono, K.; Ichijo, H.
Biochem. Biophys. Res. Commun. 253, 470-476, 1998
A:Title: Cloning and characterization of p70S6Kbeta defines a novel family of p70 S6 kin
A:Reference number: JE0377; MUID:99097259; PMID:9878560
A:Accession: JE0377
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-481 <SAI>
C:Cross-references: DDBJ:AB019245; NID:g4165310; PIDN:BA37145.1; PID:g4165311
C:Superfamily: ribosomal protein S6 kinase; protein kinase homology
C:Keywords: phosphotransferase

Query Match 50.6%; Score 43; DB 2; Length 481;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDS 12
: | | | | |
DB 381 QAFLGFTYAPS 392

RESULT 10

FtsH metalloproteinase-like protein - Arabidopsis thaliana
N:Alternate names: protein F3112.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000
C:Accession: T45642
R:Choi, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z23010

Query Match 50.6%; Score 43; DB 2; Length 481;
Best Local Similarity 58.3%; Pred. No. 32;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDS 12
: | | | | |
DB 381 QAFLGFTYAPS 392

RESULT 11

hypothetical protein Y18D10A.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26528
R:Harris, B.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20226
A:Accession: T26528
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-348 <WIL>
A:Cross-references: EMBL:AL034393; PIDN:CAA22319.1; CBSP:Y18D10A.3
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CBSP:Y18D10A.3
A:Introns: 10/2; 109/1; 168/1; 193/3

Query Match 49.4%; Score 42; DB 2; Length 348;
Best Local Similarity 58.3%; Pred. No. 33;

A:Accession: T45642
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-802 <CHO>
A:Cross-references: EMBL:AL133292
A:Experimental source: cultivar Columbia; BAC clone F13112
C:Genetics:
A:Map position: 3
A:Introns: 275/3; 296/3; 341/3; 363/3; 425/3; 454/3; 492/3; 534/2; 570/3; 612/3; 719/3;
A:Note: F13112.110

Query Match 50.6%; Score 43; DB 2; Length 802;
Best Local Similarity 60.0%; Pred. No. 56;
Matches 9; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 4 LGFSYAPPT--DSFL 16
: | | | | |
DB 624 LGFTYIPTSDRYL 638

RESULT 11

transketolase 1 VCA0624 [imported] - *Vibrio cholerae* (strain N1696) serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82437
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
I. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: C82437
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-680 <HEI>
C:Cross-references: GB:AE004392; GB:AE003853; NID:g9658031; PIDN:AAF94525.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VCA0624
A:Map position: 2
C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology

Query Match 50.0%; Score 42.5; DB 2; Length 680;
Best Local Similarity 56.2%; Pred. No. 56;
Matches 9; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

QY 3 FLGFSYAP---PTDSF 15
: | | | | |
DB 291 FLGNEYAPFEIPTDIY 306

RESULT 12

hypothetical protein Y18D10A.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26528
R:Harris, B.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20226
A:Accession: T26528
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-348 <WIL>
A:Cross-references: EMBL:AL034393; PIDN:CAA22319.1; CBSP:Y18D10A.3
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CBSP:Y18D10A.3
A:Introns: 10/2; 109/1; 168/1; 193/3

Query Match 49.4%; Score 42; DB 2; Length 348;
Best Local Similarity 58.3%; Pred. No. 33;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPT 12
: || |||| ||||
Db 240 DALFGFSHPPT 251

RESULT 13

S74753
hypothetical protein slrl610 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
S:
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74753
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-433 <KAN>

A:Cross-references: EMBL:D90901; GB:AB001339; NID:g1651897; PIDN:BAAL6904.1; PID:g165197
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: Amycolatopsis orientalis hypothetical protein PCZA361.22

Query Match 49.4%; Score 42; DB 2; Length 433;
Best Local Similarity 37.5%; Pred. No. 42;

Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
: || : ||||| : ||
Db 39 QTFDLGLGFAPPSNAYL 54

RESULT 14

S54354
inter-alpha-inhibitor H2 chain - mouse
C:Species: *Mus musculus* (house mouse)
C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999
C:Accession: S54354
R:Chan, P.; Rislter, J.L.; Raguenez, G.; Salier, J.P.
Biochem. J. 306, 505-512, 1995

A:Title: The three heavy-chain precursors for the inter-alpha-inhibitor family in mouse
A:Reference number: S54353; MUID:95194326; PMID:7534067
A:Accession: S54354
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-946 <CHA>

A:Cross-references: EMBL:X70392; NID:g695633; PIDN:CAA49842.1; PID:g695634
C:Superfamily: Inter-alpha-trypsin inhibitor complex component II

Query Match 48.8%; Score 41.5; DB 2; Length 946;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;

Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3 FLGFSYAPPTDSF 15
||| ||||| ||
Db 841 FLGI-YAPPIDKF 852

RESULT 15

H71289
hypothetical protein TP0708 - *Syphilis* spirochete
C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis* spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: H71289
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwida, R.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo-

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: H71289
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-171 <COL>

A:Cross-references: GB:AE001244; GB:AE000520; NID:g3323005; PIDN:AA05695.1; PID:g3323005
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0708

Query Match 48.2%; Score 41; DB 2; Length 171;
Best Local Similarity 58.3%; Pred. No. 23;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 LGFSYAPPTDSF 15
||:||:| ||
Db 127 LGYSFSPSTPSF 138

Search completed: August 11, 2003, 08:23:29
Job time : 4.12876 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:11:03 ; Search time 1.5794 Seconds
(without alignments)
476.401 Million cell updates/sec

Title: US-10-000-039A-4

Perfect score: 85

Sequence: 1 EAFGLFSYAPPTDSFL 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	431	1	SGK1_HUMAN
2	80	94.1	431	1	SGK1_MOUSE
3	79	92.9	430	1	SGK1_RAT
4	79	92.9	431	1	SGK1_RABIT
5	58.5	68.8	496	1	SGK3_HUMAN
6	58.5	68.8	496	1	SGK3_MOUSE
7	56	65.9	427	1	SGK2_HUMAN
8	48	56.5	367	1	SGK2_MOUSE
9	43	50.6	283	1	Y43B_MYCPN
10	43	50.6	482	1	K6B2_HUMAN
11	43	50.6	485	1	K6B2_MOUSE
12	41.5	48.8	946	1	ITH2_MOUSE
13	41	48.2	162	1	CRIZ_AGRAU
14	41	48.2	162	1	CRIZ_ALCSP
15	41	48.2	171	1	Y708_TREPA
16	41	48.2	403	1	YDHC_ECOLI
17	41	48.2	704	1	KFC1_CAEEL
18	40	47.1	272	1	Y439_MYCGE
19	40	47.1	731	1	KN3_MOUSE
20	40	47.1	732	1	KN3_RAT
21	40	47.1	952	1	UYRA_MYCGE
22	39	45.9	108	1	Y108_NPVOP
23	39	45.9	247	1	TRY2_CANFA
24	39	45.9	283	1	Y334_DEIRA
25	39	45.9	502	1	K6B1_HUMAN
26	39	45.9	502	1	K6B1_RAT
27	39	45.9	831	1	DPOF_THETH
28	38	44.7	122	1	YF93_MYCPN
29	38	44.7	246	1	YNU0_YEAST
30	38	44.7	259	1	US10_HSV4
31	38	44.7	477	1	PRP4_SCHPO
32	38	44.7	527	1	DPE2_HUMAN
33	38	44.7	533	1	ST23_HUMAN

RESULT 1

SGK1_HUMAN

ID SGK1_HUMAN STANDARD: PRT; 431 AA.

AC 000141: Q3UN56;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)

DE (Serum/glucocorticoid-regulated kinase 1).

GN SGK OR SGK1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP MEDLINE-97272242; PubMed-9114008;

RA Waldegger S., Barth P., Raber G., Lang F.;

RT "Cloning and characterization of a putative human serine/threonine

protein kinase transcriptionally modified during anisotonic and

isotonic alterations of cell volume.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:4440-4445(1997).

RN [2]

RP MEDLINE-98390195; PubMed-9722955;

RA Waldegger S., Erdel M., Nagl U.O., Barth P., Raber G., Steuer S.,

Utermann G., Paulmichl M., Lang F.;

RT "Genomic organization and chromosomal localization of the human SGK

protein kinase gene.";

RL Genomics 51:299-302(1998).

RN [3]

RP MEDLINE-22388257; PubMed-12477932;

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,

Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Richards S., Morley P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Raheij J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

SEQUENCE FROM N.A.

TISSUE=Cervix;

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,

Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Richards S., Morley P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Raheij J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

SEQUENCE FROM N.A.

TISSUE=Cervix;

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,

Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,

Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

Q920g2 mus musculus
Q99he1 neurospora
Q03563 caenorhabdi
P41003 schizosacch
Q8tcr9 homo sapien
Q59786 schizosacch
O13911 schizosacch
Q06355 solanum tub
P19823 homo sapien
P97279 mesocricetu
P37267 scharatomye
P55704 rhizobium s

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield I.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Scherch A., Schein J.B., Jones S.J.M., Maria M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-20018032; PubMed-10548550;
 RA Kobayashi T., Deak M., Morrice N., Cohen P.;
 RT "Characterization of the structure and regulation of two novel
 RT isoforms of serum- and glucocorticoid-induced protein kinase";
 RL Biochem. J. 344:189-197(1999).
 RN [6]
 RP PHOSPHORYLATION ON THR-256, AND MUTAGENESIS OF THR-256 AND SER-422.
 RX TISSUE-BRAIN;
 RX MEDLINE-99208518; PubMed-10191262;
 RA Kobayashi T., Cohen P.;
 RT "Activation of serum- and glucocorticoid-regulated protein kinase by
 RT agonists that activate phosphatidylinositol 3-kinase is mediated by
 RT 3-phosphoinositide-dependent protein kinase-1 (PDK1) and PDK2";
 RL Biochem. J. 339:319-328(1999).
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE-20345128; PubMed-10884438;
 RA Lang F., Klingel K., Wagner C.A., Stegen C., Waerntges S.,
 RA Friedrich B., Lanzendoerfer M., Melzig J., Moschen I., Steuer S.,
 RA Waldeger S., Sauter M., Paulmichl M., Gerke V., Ristler T., Gamba G.,
 RA Capasso G., Kandolf R., Hebert S.C., Massry S.G., Broer S.;
 RT "Deranged transcriptional regulation of cell-volume-sensitive kinase
 RT hSKG in diabetic nephropathy";
 RL Mol. Cell. Biol. 21:952-963(2001).
 RN [8]
 RP FUNCTION.
 RX MEDLINE-20584913; PubMed-11154281;
 RA Brunet A., Park J., Tran H., Hu L.S., Hemmings B.A., Greenberg M.E.;
 RT "Protein kinase SGK mediates survival signals by phosphorylating the
 RT forkhead transcription factor FOXO3a";
 RL Mol. Cell. Biol. 21:952-963(2001).
 RN [9]
 RP FUNCTION.
 RX MEDLINE-22284526; PubMed-12397388;
 RA Camper N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
 RA Huber S.M., Kobayashi T., Cohen P., Lang F.;
 RT "K(+)-channel activation by all three isoforms of serum- and
 RT glucocorticoid-dependent protein kinase SGK";
 RL Pfluegers Arch. 445:60-66(2002).
 CC -!- FUNCTION: Protein kinase that plays an important role in cellular
 CC stress response. Activates certain potassium, sodium, and chloride
 CC channels, suggesting an involvement in the regulation of processes
 CC such as cell survival, neuronal excitability, and renal sodium
 CC excretion. Sustained high levels and activity may contribute to
 CC conditions such as hypertension and diabetic nephropathy. Mediates
 CC cell survival signals, phosphorylates and negatively regulates
 CC pro-apoptotic FOXO3a.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 CC phosphorylation.
 CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 CC in the pancreas, followed by placenta, kidney and lung.
 CC -!- INDUCTION: By serum and/or glucocorticoids. By excessive
 CC extracellular glucose and by TGF-beta, in cultured cells.
 CC -!- PTM: Regulated by phosphorylation. Phosphoinositide 3-kinase (PI3-
 CC kinase) pathway promotes phosphorylation at Ser-422 which in turn
 CC increases the phosphorylation of Thr-256 by PDK1.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation,
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC
 DR EMBL; Y10032; CAA71138.1; ..
 DR EMBL; AJ000512; CAA04446.1; ..
 DR EMBL; AF153609; AAD41091.1; ..
 DR EMBL; BC001263; AAH01263.1; ..
 DR HSSP; P00517; LYDR.
 DR Genew; HGNC:10810; SGK.
 DR MIM; 602958; ..
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 DR GO; GO:0006950; P:response to stress; TAS.
 DR GO; GO:0006814; P:sodium ion transport; TAS.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR PROSITE; PS0107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS0108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 KW Apoptosis; Transferrase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 98 355 PROTEIN KINASE.
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1).
 FT DOMAIN 131 141 LYS/GLU-RICH.
 FT MUTAGEN 256 256 T->A: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->D: LOW ACTIVITY.
 FT MUTAGEN 256 256 T->E: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->A: LOW ACTIVITY.
 FT MUTAGEN 422 422 S->D: 10-FOLD ACTIVATION.
 FT CONFLICT 381 381 E -> D (IN REF. 3 AND 4).
 SQ SEQUENCE 431 AA; 48956 MW; F3697DAS7073399D CRC64;
 Query Match 100.0%; Score 85; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1;le-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EAFIGFSVAPPTDSFL 16
 Db 416 EAFIGFSVAPPTDSFL 431
 RESULT 2
 ID SGK1_MOUSE STANDARD; PRT; 431 AA.
 AC Q9WVC6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
 DE (Serum/glucocorticoid-regulated kinase 1).
 GN SGK OR SGK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99287894; PubMed-10148046;
 RA Naray-Rejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
 RA Rejes-Toth G.;
 RT "sgk is an aldosterone-induced kinase in the renal collecting duct.
 RT Effects on epithelial Na+ channels";
 RL J. Biol. Chem. 274:16973-16978(1999).
 RN [2]
 RP SEQUENCE FROM N.A.

RX MEDLINE=20215285; PubMed=10751222;
 RA Shigaev A., Asher C., Larter H., Garty H., Reuveny E.;
 RT "Regulation of sgk by aldosterone and its effects on the epithelial
 RT Na(+) channel.";
 RL Am. J. Physiol. 278:F613-F619(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FVB/N;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H.K., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP REGULATION BY P53.
 RC TISSUE=Mammary epithelium;
 RX MEDLINE=96218163; PubMed=8647846;
 RA Maiyar A.C., Huang A.J., Phu P.T., Cha H.H., Firestone G.L.;
 RT "p53 stimulates promoter activity of the sgk.
 RT serum/glucocorticoid-inducible serine/threonine protein kinase gene
 RT in rodent mammary epithelial cells.";
 RL J. Biol. Chem. 271:12414-12422(1996).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=22477005; PubMed=12488318;
 RA Leong M.L.L., Maiyar A.C., Kim B., O'Keefe B.A., Firestone G.L.;
 RT "Expression of the serum- and glucocorticoid-inducible protein
 RT kinase, Sgk, is a cell survival response to multiple types of
 RT environmental stress stimuli in mammary epithelial cells.";
 RL J. Biol. Chem. 278:5871-5882(2003).
 CC -!- FUNCTION: Protein kinase that plays an important role in
 CC activating certain potassium, sodium, and chloride channels,
 CC suggesting an involvement in the regulation of processes such as
 CC cell survival, neuronal excitability, and renal sodium excretion.
 CC May be a key component of cellular stress response.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 CC phosphorylation (by similarity).
 CC -!- INDUCTION: Upregulated by tumor suppressor p53 in mammary
 CC epithelial tumor cells.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: AF139638; AAD43302.1;
 CC EMBL: AF205855; AAF19429.1;
 CC EMBL: BC005720; AAF05720.1;
 CC HSP: P00517; 1YDR.
 CC MGD: MGI:1340062; Sgk.
 CC InterPro: IPR000961; Pkinase_C.
 CC InterPro: IPR000719; Prot_kinase.

DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00433; pkinase_C; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW Apoptosis; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 98 355 PROTEIN KINASE.
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 137 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1) (BY
 FT SIMILARITY).
 SQ SEQUENCE 431 AA; 48928 MW; 6DF5B846AAC2754 CRC64;

Query Match 94.1%; Score 80; DB 1; Length 431;
 Best Local Similarity 93.8%; Pred No. 7.2e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16

|||||

DB 416 EAFLGFSYAPPTDSFL 431

RESULT 3

SGK1_RAT STANDARD; PRT; 430 AA.

ID SGK1_RAT

AC Q06226;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Serine/threonine-protein kinase Sgk1 (p53 2.1.37)

DE (Serum/glucocorticoid-regulated kinase 1).

GN SGK OR SGK1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Fischer 344;

RX MEDLINE=93204949; PubMed=8455596;

RA Webster M.K., Goya L., Ge Y., Maiyar A.C., Firestone G.L.;

RT "Characterization of sgk, a novel member of the serine/threonine

RT protein kinase gene family which is transcriptionally induced by

RT glucocorticoids and serum.";

RL Mol. Cell. Biol. 13:2031-2040(1993).

RN [2]

RP INDUCTION BY CNS INJURY.

RX MEDLINE=95151713; PubMed=7854047;

RA Imaizumi K., Tsuda M., Wanka A., Tohyama M., Takagi T.;

RT "Differential expression of sgk mRNA, a member of the Ser/Thr protein

RT kinase gene family, in rat brain after CNS injury.";

RL Brain Res. Mol. Brain Res. 26:189-196(1994).

RN [3]

RP INDUCTION BY FSH.

RX MEDLINE=95258633; PubMed=7740159;

RA Richards J.S., Fitzpatrick S.L., Clemens J.W., Morris J.K.,

RA Alliston T., Siros J.;

RT "Ovarian cell differentiation: a cascade of multiple hormones,

RT cellular signals, and regulated genes.";

RL Recent Prog. Horm. Res. 50:223-254(1995).

RN [4]

RP INDUCTION BY P53.

RC TISSUE=Mammary epithelium;

RX MEDLINE=96218163; PubMed=8647846;

RA Maiyar A.C., Huang A.J., Phu P.T., Cha H.H., Firestone G.L.;

RT "p53 stimulates promoter activity of the sgk.

RT serum/glucocorticoid-inducible serine/threonine protein kinase gene

in rodent mammary epithelial cells.";
 J. Biol. Chem. 271:12414-12422(1996).
 [5]
 PHOSPHORYLATION ON THR-256 BY PDPK1.
 MEDLINE-99286226; PubMed-10357815;
 Park J., Leong M.L., Buse P., Malyar A.C., Firestone G.L.,
 Hemmings B.A.;
 "Serum and glucocorticoid-inducible kinase (SGK) is a target of the PI
 3-kinase-stimulated signaling pathway.";
 EMBO J. 18:3024-3033(1999).
 -!- FUNCTION: Protein kinase that plays an important role in
 activating certain potassium, sodium, and chloride channels,
 suggesting an involvement in the regulation of processes such as
 cell survival, neuronal excitability, and renal sodium excretion.
 May be a key component of cellular stress response (By
 similarity). May also play an important role in the development of
 particular groups of neurons in the postnatal brain.
 -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 phosphorylation (By similarity).
 -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
 in the ovary, thymus and lung.
 -!- INDUCTION: By dexamethasone and serum. By tumor suppressor p53 in
 mammary epithelial tumor cells. By FSH in granulosa cells. By
 injury to the central nervous system.
 -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; L01624; AAA2137.1; --
 DR HSP; P00517; IYDR.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR Apoptosis; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW phosphorylation.
 FT DOMAIN 98 354 PROTEIN KINASE.
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1).
 SQ SEQUENCE 430 AA; 48927 MW; OD5845B04156F26D CRC64;

 Query Match 92.9%; Score 79; DB 1; Length 430;
 Best Local Similarity 93.8%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 415 EAFLGFSYAPPTDSFL 430

 RESULT 4
 SGK1_RABIT
 ID SGK1_RABIT STANDARD; PRT; 431 AA.
 AC Q9XT18;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

Serine/threonine-protein kinase Sgk1 (EC 2.7.1.37)
 (Serum/glucocorticoid-regulated kinase 1).
 SGK OR SGK1.
 Oryctolagus cuniculus (Rabbit).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NCBI_TaxID=9986;
 [1]
 SEQUENCE FROM N.A.
 STRAIN-New Zealand white;
 MEDLINE-99287894; PubMed-10358046;
 Naray-Fejes-Toth A., Canessa C., Cleaveland E.S., Aldrich G.,
 Naray-Fejes-Toth G.;
 "Sgk1 is an aldosterone-induced kinase in the renal collecting duct".
 Effects on epithelial Na+ channels.";
 J. Biol. Chem. 274:16973-16978(1999).
 -!- FUNCTION: Protein kinase that plays an important role in
 activating certain potassium, sodium, and chloride channels,
 suggesting an involvement in the regulation of processes such as
 cell survival, neuronal excitability, and renal sodium excretion.
 May be a key component of cellular stress response (By
 similarity).
 -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear, upon
 phosphorylation (By similarity).
 -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; AF139639; AAD43303.1; --
 DR HSP; P00517; IYDR.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_ST; 1.
 DR Apoptosis; Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW phosphorylation.
 FT DOMAIN 98 355 PROTEIN KINASE.
 FT NP_BIND 104 112 ATP (BY SIMILARITY).
 FT BINDING 127 127 ATP (BY SIMILARITY).
 FT ACT_SITE 222 222 BY SIMILARITY.
 FT MOD_RES 256 256 PHOSPHORYLATION (BY PDPK1) (BY
 SIMILARITY).
 SQ SEQUENCE 431 AA; 48999 MW; 354898A77E8E38FD CRC64;

 Query Match 92.9%; Score 79; DB 1; Length 431;
 Best Local Similarity 93.8%; Pred. No. 1e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 416 EAFLGFSYAPPTDSFL 431

 RESULT 5
 SGK3_HUMAN
 ID SGK3_HUMAN STANDARD; PRT; 496 AA.
 AC Q9GBR1; Q9P107; Q9UKG5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)
Serine/threonine-protein kinase Sgk3 (EC 2.7.1.37)
(Serum/glucocorticoid regulated kinase 3) (Serum/glucocorticoid
regulated kinase-like).
SGK3 OR SGK3.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBTaxID=9606;
[1]
SEQUENCE FROM N.A., PHOSPHORYLATION ON THR-320, AND MUTAGENESIS OF
SER-486.
RP MEDLINE-20018032; PubMed-10548550;
RX Kobayashi T., Deak M., Morrice N., Cohen P.;
RA "Characterization of the structure and regulation of two novel
RT isoforms of serum- and glucocorticoid-induced protein kinase.";
KL Biochem. J. 344:189-197(1999).
[2]
SEQUENCE FROM N.A.
RP MEDLINE-20054360; PubMed-10585774;
RX Dai F., Yu L., He H., Zhao Y., Yang J., Zhang X., Zhao S.;
RA "Cloning and mapping of a novel human Serum/Glucocorticoid regulated
RT kinase-like gene, SGK3, to chromosome 8q12.3-q13.1.";
KL Genomics 62:95-97(1999).
[3]
REVIEWS.
RA Zhao Y.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RP TISSUE-Breast;
RX MEDLINE-22388257; PubMed-12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
FUNCTION.
RP MEDLINE-22284526; PubMed-12397388;
RX Gamber N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
RA Huber S.M., Kobayashi T., Cohen P., Lang F.;
RT "K(+) channel activation by all three isoforms of serum- and
glucocorticoid-dependent protein kinase SGK.";
PLfugers Arch. 445:60-66(2002).
RL -!- FUNCTION: Involved in the activation of potassium channels.
CC -!- CATALYTIC ACTIVITY: ATP + a protein - ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Localized in vesicle-like structures and in
the early endosome (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in most tissues with highest levels
in pancreas, kidney liver, heart and brain and lower levels in
lung, placenta and skeletal muscle.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation.

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF169035; AAF12758.1; ALT_INIT.
DR EMBL; AF085233; AAF27051.2; -.
DR EMBL; BC015326; AAH15326.1; -.
DR Genew; HGNC:10812; SGK3.
DR MIM; 607591; -.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001683; PX.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00433; Pkinase_C; 1.
DR Pfam; PF00787; PX; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00195; PX; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 12 124 PX.
FT DOMAIN 162 419 PROTEIN KINASE.
FT NP_BIND 168 176 ATP (BY SIMILARITY).
FT BINDING 191 191 ATP (BY SIMILARITY).
FT ACT_SITE 286 286 BY SIMILARITY.
FT MOD_RES 320 320 PHOSPHORYLATION (BY PDPK1).
FT MUTAGEN 486 486 S->D: INCREASED ACTIVATION.
FT CONFLICT 54 54 F -> V (IN REF. 2 AND 3).
FT CONFLICT 127 127 D -> G (IN REF. 2 AND 3).
FT CONFLICT 187 187 F -> V (IN REF. 2 AND 3).
FT CONFLICT 294 294 L -> V (IN REF. 2 AND 3).
FT CONFLICT 387 387 T -> R (IN REF. 2 AND 3).
SQ SEQUENCE 496 AA; 57108 MW; 76A6CCEB69006CF1 CRC64;
Query Match 68.88; Score 58.5; DB 1; Length 496;
Best Local Similarity 70.68; Pred. No. 0.032; Indels 1; Gaps 1;
Matches 12; Conservative 3; Mismatches 1;
Qy 1 EAFVGFSAAPPT-DSFL 16
:|||||:|
Db 480 DAFVGFSAAPSEDLFL 496
RESULT 6
SGK3_MOUSE STANDARD; PRT; 496 AA.
ID SGK3_MOUSE
AC Q9ER3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Sgk3 (EC 2.7.1.37)
DE (Serum/glucocorticoid regulated kinase 3) (Serum/glucocorticoid
DE regulated kinase-like) (Cytokine independent survival kinase).
DE SGK3 OR SGK3 OR CISK.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBTaxID=10090;
RN
RP SEQUENCE FROM N.A. (ISOFORM 1), AND MUTAGENESIS OF LYS-191.
RX MEDLINE-20504817; PubMed-11050396;
RA Liu D., Yang X., Songyang Z.;
RT "Identification of CISK, a new member of the SGK kinase family that
RT promotes IL-3-dependent survival.";
RL Curr. Biol. 10:1233-1236(2000).

[2] SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP STRAIN-C57BL/6J; TISSUE=Forelimb, Ovary, and uterus;
 RC MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Futano M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gofjohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schröml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Graestland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maitalis L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numa K., Okido T., Pavan W.J., Perteira G., Resole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wallstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.;
 RL Nature 420:563-573(2002).
 RN [3]
 RP CHARACTERIZATION, AND MUTAGENESIS OF ARG-90.
 RX MEDLINE=21405768; PubMed=11514587;
 RA Xu J., Liu D., Gill G., Songyang Z.;
 RT Regulation of cytokine-independent survival kinase (CISK) by the Phox
 RT homology domain and phosphoinositides.;
 RL J. Cell Biol. 154:695-705(2001).
 CC -!- FUNCTION: Involved in the activation of potassium channels (By
 CC similarity). Mediates cell IL-3-dependent survival signals. Can
 CC inhibit pro-apoptotic FOXO3A in vitro.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Localized in vesicle-like structures and in
 CC the early endosome.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-alternative splicing; Named isoforms=2;
 CC Name=1;
 CC isoId=Q9ERE3-1; Sequence=Displayed;
 CC Name=2;
 CC isoId=Q9ERE3-2; Sequence=VSP_004934, VSP_004935;
 CC Note-No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Widely expressed, predominantly in the heart,
 CC spleen and 7-day embryo.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 1 phox homology (PX) domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF312007; AAC34115.1; -;
 CC EMBL: AK028722; BAC26083.1; -;
 CC EMBL: AK030314; BAC26895.1; -;
 CC EMBL: AK031133; BAC27269.1; -;
 CC EMBL: AK031328; BAC27349.1; -;
 CC HSP: P05132; IATP.
 CC MGD: MGI:2182368; Sgk3.
 DR GO: 0016023; Cytoplasmic vesicle; IDA.
 DR GO: 0008189; F-actin polymerization; IDA.
 DR GO: 0004674; F-actin polymerization; IDA.
 DR GO: 0006916; F-actin polymerization; IDA.
 DR InterPro: IPR000961; pkinase_C.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR001683; PX.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00433; pkinase_1.
 DR Pfam: PF00787; PX.
 DR ProDom: PD000001; Prot_kinase_1.
 DR SMART: SM00312; PX.
 DR SMART: SM00313; S_TK_X.
 DR SMART: SM00220; S_TK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00193; PX; 1.
 DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
 KW phosphorylation; Alternative splicing.
 FT DOMAIN 12 124
 FT NP_BIND 162 419
 FT BINDING 168 176
 FT ACT_SITE 191 191
 FT MOD_RES 286 286
 FT VARSPPLIC 327 333
 FT VARSPPLIC 334 496
 FT MUTAGEN 90 90
 FT MUTAGEN 191 191
 FT CONFLICT 114 114
 FT CONFLICT 204 204
 FT SEQUENCE 496 AA; 57145 MW; 4B7D2804A5948BAD CRC64;
 Query Match 68.8%; Score 58.5; DB 1; Length 496;
 Best Local Similarity 70.6%; Pred. No. 0.032; 1; Indels 1; Gaps 1;
 Matches 12; Conservative 3; Mismatches 1;
 QY 1 EAFICFSVAPPT-DSFL 16
 DB 480 DAFVGFSTAPSEDLFL 496
 RESULT 7
 SGK2_HUMAN
 ID SGK2_HUMAN STANDARD; PRT; 427 AA.
 AC Q9H8Y8; Q9H8Y8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Sgk2 (EC 2.7.1.37)
 DE (Serum/glucocorticoid regulated kinase 2).
 GN SGK2
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1 AND 2), PHOSPHORYLATION ON THR-253, AND
 RP MUTAGENESIS OF SER-416.
 RX MEDLINE=20018032; PubMed=10548550;
 RA Kobayashi T., Deak M., Morrice N., Cohen P.;
 RT "Characterization of the structure and regulation of two novel
 RL isoforms of serum- and glucocorticoid-induced protein kinase.";
 RN Biochem. J. 344:189-197(1999).
 RP [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=21638749; PubMed=11780052;

RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights K., Laird G.K., Lawlor S.,
RA Leharvaish M.H., Leversha M.A., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLean K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.L.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
[3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE-Colon:
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Aitschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshlyuki S., Carninci P., Frange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.G., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RP FUNCTION.
RX MEDLINE=22284526; PubMed=12397388;
RA Gamber N., Fillon S., Feng Y., Friedrich B., Lang P.A., Henke G.,
RA Huber S.M., Kobayashi T., Cohen P., Lang F.;
RT "K(+) channel activation by all three isoforms of serum- and
glucocorticoid-dependent protein kinase SGK.";
PLugers Arch. 445:60-66(2002).
RL -!- FUNCTION: Involved in the activation of potassium channels.
CC -!- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=2; Synonyms=beta;
CC IsoId-Q9HBY8-1; Sequence=Displayed;
CC Name=1; Synonyms=alpha;
CC IsoId-Q9HBY8-2; Sequence=VSP_004932;
CC -!- TISSUE SPECIFICITY: Highly expressed in liver, kidney and
pancreas, and at lower levels in brain.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- CAUTION: Not regulated by serum or glucocorticoids.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF186470; AAG17012.1; -
DR EMBL: AF169034; AAF12757.2; -
DR EMBL: Z98752; CAC18509.1; -
DR EMBL: BC014037; AAH14037.1; -
DR HSSP: P05132; IATP.
DR Genew; HGNC:13900; SGK2.
DR MIM; 607589; -
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; prot_kinase.
DR Pfam; PF00089; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; prot_kinase; 1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_ST; 1.
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 95 352 PROTEIN KINASE.
FT NP_BIND 101 109 ATP (BY SIMILARITY).
FT BINDING 124 124 ATP (BY SIMILARITY).
FT ACT_SITE 219 219 BY SIMILARITY.
FT MOD_RES 253 253 PHOSPHORYLATION (BY PDPK1).
FT VARSPPLIC 1 60 Missing (in isoform 1).
FT FTID-VSP_004932.
FT MUTAGEN 416 416 S->D: INCREASED ACTIVATION.
SQ SEQUENCE 427 AA; 47604 MW; D8F0FA6DF54B1370 CRC64;
Query Match 65.9%; Score 56; DB 1; Length 427;
Best Local Similarity 73.3%; Pred. No. 0.071;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AFLGFSVAPPTDSFL 16
DB 411 AFLGFSAPEDDDIL 425
RESULT 8
SGK2_MOUSE STANDARD; PRT; 367 AA.
AC Q9QZS5; Q8R0P6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase SGK2 (EC 2.7.1.-)
DE (Serum/glucocorticoid regulated kinase 2).
GN SGK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20018032; PubMed=10548550;
RA Kobayashi T., Deak M., Morrice N., Cohen P.;
RT "Characterization of the structure and regulation of two novel
isoforms of serum- and glucocorticoid-induced protein kinase.";
RL Biochem. J. 344:189-197(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

```

DR EMBL: AK050009; BAC34031.1; -
DR EMBL: BC026549; AAH26549.1; -
DR HSSP: P05132; 1CTP.
DR MGD: MGI:1351318; Sgk2.
DR InterPro: IPR000961; pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase_C; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00133; S_TK_X; 1.
DR SMART: SM0220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR KX Transfaser; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 35 292
FT NP_BIND 41 49 ATP (BY SIMILARITY).
FT BINDING 64 64 ATP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
FT MOD_RES 193 193 PHOSPHORYLATION (BY PDPK1)
FT FT (BY SIMILARITY).
FT VARSPIC 171 199 Missing (in isoform 2).
FT CONFLICT 77 77 /FTID:VSP_004933.
FT CONFLICT 77 77 MISSING (IN REF. 3).
SQ SEQUENCE 367 AA; 61359 MW; 668C04B1A1E9E33A CRC64;

Query Match 56.5%; Score 48; DB 1; Length 367;
Best local Similarity 66.7%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AFLGFSYAPPTDSFL 16
DB 351 AFLGFSYAQDDDL 365
||||||| 1 1
||||||| 1 1

RESULT 9
Y43B_MYCPN STANDARD; PRT; 283 AA.
ID Y43B_MYCPN
AC P75152;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical lipoprotein MG439 homolog 2 precursor (E09_orf283a).
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
CN NCBI_Taxid=2104;
FN 11
PP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
EX MEDLINE=97105885; Pubmed=8948633;
RA Himmelfreuch R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
FT "Complete sequence analysis of the genome of the bacterium Mycoplasma
FT pneumoniae."
FT Nucleic Acids Res. 24:4420-4449(1996).
RL CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE MG439 / MG440 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000020; AAB95845.1; -
CC PIR: S73523; S73523.
DR InterPro: IPR001595; Lipoprotein_3.

```

DR Pfam: PF00938; Lipoprotein_3; 1.
 DR ProDom: PD003276; Lipoprotein_3; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Hypothetical protein; Lipoprotein; Membrane; Signal;
 KW Complete proteome.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 283 HYPOTHETICAL LIPOPROTEIN MG439 HOMOLOG 2.
 FT LIPID 22 22 N-ACYL DIGLYCERIDE (POTENTIAL).
 SQ SEQUENCE 283 AA; 31355 MW; 3A55F5B8B55B005 CRC64;
 Query Match 50.6%; Score 43; DB 1; Length 283;
 Best Local Similarity 50.0%; Pred. No. 6.7;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPTDS 14
 Db 204 DGFTFTYPTQTS 217
 RESULT 10
 ID_K6B2_HUMAN STANDARD; PRT; 482 AA.
 AC Q9UBS0; Q94809; Q9UEC1;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ribosomal protein S6 kinase beta 2 (EC 2.7.1.1) (S6K-beta 2) (70 kDa
 DE ribosomal protein S6 kinase 2) (p70-S6KB) (p70 ribosomal S6 kinase
 DE beta) (p70 S6Kbeta) (S6K2) (S6 kinase-related kinase) (SRK)
 DE (Serine/threonine-protein kinase 14 beta).
 GN RPS6KB2 OR STK14B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99023916; PubMed=9804755;
 RA Gout I., Minami T., Hara K., Tsuchishita Y., Filonenko V.,
 RA Waterfield M.D., Yonezawa K.;
 RT "Molecular cloning and characterization of a novel p70 S6 kinase, p70
 RT S6 kinase beta containing a proline-rich region.";
 RL J. Biol. Chem. 273:30061-30064(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99097259; PubMed=9878560;
 RA Saitoh M., ten Dijke P., Miyazono K., Ichijo H.;
 RT "Cloning and characterization of p70 S6Kbeta defines a novel family of
 RT p70 S6 kinases.";
 RL Biochem. Biophys. Res. Commun. 253:470-476(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99422045; PubMed=10490847;
 RA Lee-Fruman K.K., Kuo C.J., Lippincott J., Terada N., Blenis J.;
 RT "Characterization of S6K2, a novel kinase homologous to S6K1.";
 RL Oncogene 18:5108-5114(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Koh H.J., Lee B.N., Choi H.S., Chung J.;
 RT "Cloning and characterization of a novel S6 kinase-related kinase,
 RT SRK.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan T., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.R.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: PHOSPHORYLATES SPECIFICALLY RIBOSOMAL PROTEIN S6.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC S6 KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB016869; BAA34402.1; ALT_INIT.
 CC EMBL; AB019245; BAA37145.1; -
 CC EMBL; AF076931; AAD46063.1; -
 CC EMBL; AF099739; AAD20990.1; -
 CC EMBL; BC000094; AHH00094.1; -
 CC PIR; JE0377; JE0377.
 CC HSP; Q63450; IAO6.
 CC Genew; HGNC:10437; RPS6KB2.
 CC GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 CC GO; GO:0006412; P:protein biosynthesis; TAS.
 CC GO; GO:0000074; P:regulation of cell cycle; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000961; Prot_kinase_C.
 CC InterPro; IPR00719; Prot_kinase.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF00433; pkinase_C; 1.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00133; S_TK_X; 1.
 CC SMART; SM00220; S_TKC; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; Phosphorylation;
 KW ATP-binding.
 FT DOMAIN 67 328 PROTEIN KINASE.
 FT DOMAIN 411 482 PRO-RICH.
 FT NP_BIND 73 81 ATP (BY SIMILARITY).
 FT BINDING 99 99 ATP (BY SIMILARITY).
 FT ACT_SITE 194 194 BY SIMILARITY.
 FT CONFLICT 90 90 MISSING (IN REF. 2).
 FT CONFLICT 409 409 R -> C (IN REF. 2).
 FT CONFLICT 420 420 V -> A (IN REF. 2).
 SQ SEQUENCE 482 AA; 53483 MW; 5D2C691D24A3D0A9 CRC64;
 Query Match 50.6%; Score 43; DB 1; Length 482;
 Best Local Similarity 58.3%; Pred. No. 12;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EAFLGFSYAPPT 12
 Db 382 QAFGLGFTYVAPS 393
 RESULT 11
 ID_K6B2_MOUSE STANDARD; PRT; 485 AA.
 AC Q9Z1M4;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Ribosomal protein S6 kinase beta 2 (EC 2.7.1.-) (S6K-beta 2) (70 kDa
 DE ribosomal protein S6 kinase 2) (p70-S6KB) (p70 ribosomal S6 kinase
 DE beta) (p70 S6Kbeta) (S6K2).
 GN RPS6KB2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=9822608; PubMed=9822608;
 RX Shima H., Pende M., Chen Y., Funagalli S., Thomas G., Kozma S.C.;
 RA "Disruption of the p70(S6K)/p85(S6K) gene reveals a small mouse
 RT phenotype and a new functional S6 kinase";
 RL EMBO J. 17:6649-6659(1998).
 CC -!- FUNCTION: PHOSPHORYLATES SPECIFICALLY RIBOSOMAL PROTEIN S6.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC S6 KINASE SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AJ007938; CAA07774.1;
 DR HSSP: Q63450; IAO6.
 DR MGI: I927343; Rps6kb2.
 DR InterPro: IPR000961; pkinase_C.
 DR InterPro: IPR000719; prot_kinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00693; pkinase_1.
 DR Pfam: PF00433; pkinase_C_1.
 DR ProDom: PD000001; Prot_kinase: 1.
 DR SMART: SM00133; S_TK_X; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR Transferrase; Serine/threonine-protein kinase; Phosphorylation;
 KW ATP-binding. 67 328 PROTEIN KINASE.
 FT DOMAIN 411 485 PRO-RICH.
 FT NP_BIND 73 81 ATP (BY SIMILARITY).
 FT BINDING 99 99 ATP (BY SIMILARITY).
 FT ACT_SITE 194 194 BY SIMILARITY.
 SQ SEQUENCE 485 AA; 53538 MW; 39629ADAR0F6C66 CRC64;
 Query Match 50.68; Score 43; DB 1; Length 485;
 Best Local Similarity 58.38; Pred. No. 12; Mismatches 0; Gaps 0;
 Matches 7; Conservative 3; Indels 2;
 QY 1 EAFLEFSYAPPT 12
 DB 382 QAFLEFTYVAPS 393
 RESULT 12
 ITH2_MOUSE
 ID ITH2_MOUSE STANDARD; PRT; 946 AA.
 AC Q61703;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITH1 heavy
 DE chain H2) (Inter-alpha-inhibitor heavy chain 2).
 GN ITH2.
 OS Mus musculus (Mouse).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6N; TISSUE=Liver;
 RX MEDLINE=95194326; PubMed=7534067;
 RA Chan P., Ristler J.L., Raguenez G., Sallier J.-P.;
 RT "The three heavy-chain precursors for the inter-alpha-inhibitor
 RT family in mouse: new members of the multicopper oxidase protein group
 RT with differential transcription in liver and brain";
 RL Biochem. J. 306:505-512(1995).
 CC -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
 CC BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
 CC INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
 CC LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
 CC ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
 CC SIMILARITY).
 CC -!- SUBUNIT: I-ALPHA-1 PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
 CC ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
 CC BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-1) IS COMPOSED OF H1, H2
 CC AND BIKUNIN. INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
 CC BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH LIVER AND BRAIN.
 CC -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
 CC 4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE ITH FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 VWFA DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X70392; CAA49842.1;
 DR PIR: S54354; S54354.
 DR MGI: 96619; Ith2.
 DR InterPro: IPR006587; VIT.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF00092; vwa_1.
 DR SMART: SM00609; VIT; 1.
 DR SMART: SM00327; VWFA; 1.
 DR PROSITE: PS50234; VWFA; 1.
 DR Serine protease inhibitor; Repeat; Signal; Multigene family;
 KW Glycoprotein.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 54 BY SIMILARITY.
 FT CHAIN 55 702 INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
 FT H2.
 FT PROPEP 703 946 BY SIMILARITY.
 FT DOMAIN 308 468 VWFA.
 FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT BINDING 702 702 CHONDROITIN 4-SULFATE, CROSS-LINK SITE
 FT (BY SIMILARITY).
 SQ SEQUENCE 946 AA; 105927 MW; 40DB6716433ED9DC CRC64;
 Query Match 48.88; Score 41.5; DB 1; Length 946;
 Best Local Similarity 69.28; Pred. No. 42;
 Matches 9; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 3 FLGFSYAPPTDSF 15
 DB 841 FLGI-YAPPIDKF 852
 RESULT 13
 CRTZ_AGRAU

```

ID CRTZ_AGRAU STANDARD; PRT: 162 AA.
AC P54973;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta-carotene hydroxylase.
GN CRTZ.
OS Agrobacterium aurantiacum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=44155;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96082243; PubMed=7592436;
RA Misawa N., Satomi Y., Kondo K., Yokoyama A., Kajiwara S., Saito T.,
RA Ohtani T., Miki W.;
RT "Structure and functional analysis of a marine bacterial carotenoid
RT biosynthesis gene cluster and astaxanthin biosynthetic pathway
RT proposed at the gene level."
RL J. Bacteriol. 177:6575-6584(1995).
CC -!- FUNCTION: Catalyzes the hydroxylation reaction from
CC beta-carotene to zeaxanthin via beta-cryptoxanthin.
CC -!- PATHWAY: CAROTENOID BIOSYNTHESIS. INVOLVED IN ASTAXANTHIN
CC BIOSYNTHETIC PATHWAY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D58420; BAA09592.1;
DR InterPro: IPR005596; Carotene_hydrox.
DR Pfam: PF03897; Carotene_hydrox; 1.
KW Carotenoid biosynthesis.
SQ SEQUENCE 162 AA; 18926 MW; FB8891AA12A92C6C CRC64;

Query Match 48.2%; Score 41; DB 1; Length 162;
Best Local Similarity 77.8%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GFSYAPPTD 13
DB 134 GFIVAPPVD 142

RESULT 14
CRTZ_ALCSP STANDARD; PRT: 162 AA.
AC Q44262;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta-carotene hydroxylase.
GN CRTZ.
OS Alcaligenes sp.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Alcaligenes.
OX NCBI_TaxID=512;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95251715; PubMed=7733978;
RA Misawa N., Kajiwara S., Kondo K., Yokoyama A., Satomi Y., Saito T.,
RA Miki W., Ohtani T.;
RT "Canthaxanthin biosynthesis by the conversion of methylene to keto
RT groups in a hydrocarbon beta-carotene by a single gene."
RL Biochem. Biophys. Res. Commun. 209:867-876(1995).
CC -!- FUNCTION: Catalyzes the hydroxylation reaction from beta-carotene
CC to zeaxanthin via beta-cryptoxanthin (By similarity).
CC -!- PATHWAY: CAROTENOID BIOSYNTHESIS. INVOLVED IN ASTAXANTHIN

```

```

CC BIOSYNTHETIC PATHWAY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D58422; BAA09597.1;
DR InterPro: IPR005596; Carotene_hydrox.
DR Pfam: PF03897; Carotene_hydrox; 1.
KW Carotenoid biosynthesis.
SQ SEQUENCE 162 AA; 19157 MW; QCA5235CBFB33B5 CRC64;

Query Match 48.2%; Score 41; DB 1; Length 162;
Best Local Similarity 77.8%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 GFSYAPPTD 13
DB 134 GFIVAPPVD 142

RESULT 15
Y708_TREPA STANDARD; PRT: 171 AA.
ID Y708_TREPA
AC O83706;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0708.
GN TP0708.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AE001244; AAC65695.1;
DR PIR: H71289; H71289.
DR TIGR: TP0708;
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 13 35 POTENTIAL.
FT TRANSMEM 50 72 POTENTIAL.
SQ SEQUENCE 171 AA; 17961 MW; 8C97AB403F9A3384 CRC64;

Query Match 48.2%; Score 41; DB 1; Length 171;
Best Local Similarity 58.3%; Pred. No. 8.5;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

Qy 4 LGFSYAPPTDSF 15
Db 127 LGYSFSPSTSE 138

Search completed: August 11, 2003, 08:19:56
Job time : 3.5794 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 07:58:29 ; Search time 4.29185 Seconds
(without alignments)
962.020 Million cell updates/sec

Title: US-10-000-039A-4
Perfect score: 85
Sequence: 1 EAFLGFSYAPPTDPSFL 16

Scoring Table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPREMBL_23.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-rvivirus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	94.1	185	11 Q99LU4	Q99LU4 mus musculus
2	76	89.4	433	13 Q73926	Q73926 squalus aca
3	76	89.4	594	13 Q73927	Q73927 squalus aca
4	72	84.7	40	6 Q95N98	Q95N98 canis famill
5	70	82.4	434	13 Q95524	Q95524 xenopus lae
6	58.5	68.8	429	11 Q8VEK1	Q8VEK1 mus musculus
7	54	63.5	374	17 Q97YG2	Q97YG2 sulfolobus
8	54	63.5	374	17 Q97X02	Q97X02 sulfolobus
9	48	56.5	375	17 Q972T2	Q972T2 sulfolobus
10	45.5	53.5	164	16 Q8RH02	Q8RH02 fusbacteri
11	45	52.9	330	16 Q8X834	Q8X834 escherichia
12	45	52.9	341	5 Q8MKW0	Q8MKW0 drosophila
13	45	52.9	411	5 Q9V5E5	Q9V5E5 drosophila
14	45	52.9	896	10 Q93YS3	Q93YS3 arabidopsis
15	45	52.9	910	10 Q9C520	Q9C520 arabidopsis
16	44	51.8	390	17 Q82T17	Q82T17 pyrobaculum

17	44	51.8	516	16 Q8P3U9	Q8P3U9 xanthomonas
18	44	51.8	533	16 Q8PFC8	Q8PFC8 xanthomonas
19	44	51.8	592	13 Q9PVD1	Q9PVD1 xenopus lae
20	43	50.6	162	3 O14190	O14190 schizosacch
21	43	50.6	208	11 Q99JU0	Q99JU0 mus musculus
22	43	50.6	326	11 Q91V80	Q91V80 mus musculus
23	43	50.6	359	16 Q8R863	Q8R863 thermocae
24	43	50.6	444	16 Q9A5G5	Q9A5G5 caulobacter
25	43	50.6	482	4 Q9BR50	Q9BR50 homo sapien
26	43	50.6	802	10 Q9SD67	Q9SD67 arabidopsis
27	43	50.6	806	10 Q9FIM2	Q9FIM2 arabidopsis
28	43	50.6	2426	3 Q96U0R	Q96U0R aspergillus
29	42.5	50.0	680	16 Q9K1W7	Q9K1W7 vibrio chol
30	42	49.4	219	16 Q9CK02	Q9CK02 pasteurilla
31	42	49.4	348	5 Q9XW15	Q9XW15 caenorhabdi
32	42	49.4	385	2 Q9X4F1	Q9X4F1 rhodobacter
33	42	49.4	433	16 P72888	P72888 syntrophocyst
34	42	49.4	743	5 Q8MWS3	Q8MWS3 babesia bov
35	42	49.4	857	10 Q9SB20	Q9SB20 phaseolus a
36	42	49.4	3151	5 Q8SR52	Q8SR52 encephalito
37	41.5	48.8	946	11 Q8K016	Q8K016 mus musculus
38	41.5	48.8	1248	5 Q9NGK5	Q9NGK5 drosophila
39	41	48.2	162	2 Q9RLH6	Q9RLH6 paracoccus
40	41	48.2	166	11 Q8C7P6	Q8C7P6 mus musculus
41	41	48.2	181	11 Q8BPL2	Q8BPL2 mus musculus
42	41	48.2	197	4 Q8N8D7	Q8N8D7 homo sapien
43	41	48.2	277	9 Q94MX5	Q94MX5 enterobacte
44	41	48.2	299	5 Q8IT95	Q8IT95 caenorhabdi
45	41	48.2	335	16 Q9RTQ3	Q9RTQ3 deinorococcus

ALIGNMENTS

RESULT 1
Q99LU4 PRELIMINARY: PRT; 185 AA.
ID Q99LU4;
AC Q99LU4;
DC 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to serum/glucocorticoid regulated kinase (Fragment).
GN SGK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002222; AAH02222.1;
DR MGD; MGI:1340062; Sgk.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD0000001; Prot_kinase; 1.
DR SMART; SM00133; S-TK_X; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 185 AA: 20921 MW; EDAA44ABF083945 CRC64;

Query Match 94.1%; Score 80; DB 11; Length 185;
Best Local Similarity 93.8%; Pred. NO. 5.4e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDPSFL 16
|||||
Db 170 EAFLGFSYAPPTDPSFL 185


```

RESULT 2
073926 PRELIMINARY: PRT; 433 AA.
AC 073926:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE S-sgk1.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
SEQUENCE FROM N.A.
RP Waldeyer S., Barth P., Forrest J., Greger R.F., Lang F.,
RT "Cloning of sgk Serine-Threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AJ223715; CAA11527.1; -.
DR HSSP: P05132; IFMO.
DR InterPro: IPR000961; pkinase.C.
DR InterPro: IPR000719; prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR Pfam: PF00433; pkinase.C.1.
DR ProDom: PD000001; prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00133; S_TK_X; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 433 AA; 48369 MW; 982BD01A59AA3A78 CRC64;

Query Match 89.4%; Score 76; DB 13; Length 433;
Best Local Similarity 87.5%; Pred. No. 6.7e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
:|||||
DB 418 DAFLGFSYAPPMDSFL 433

RESULT 3
073927 PRELIMINARY: PRT; 594 AA.
AC 073927:
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE S-sgk2.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
SEQUENCE FROM N.A.
RP Waldeyer S., Barth P., Forrest J., Greger R.F., Lang F.,
RT "Cloning of sgk Serine-Threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AJ223716; CAA11528.1; -.
DR HSSP: P00517; 1VDR.
DR InterPro: IPR000961; pkinase.C.
DR InterPro: IPR000719; prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase.C.1.
DR ProDom: PD000001; prot_kinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00133; S_TK_X; 1.

```

```

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 594 AA; 66857 MW; 40153EC18D57B767 CRC64;

Query Match 89.4%; Score 76; DB 13; Length 594;
Best Local Similarity 87.5%; Pred. No. 9.4e-05;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
:|||||
DB 579 DAFLGFSYAPPMDSFL 594

RESULT 4
095N98 PRELIMINARY: PRT; 40 AA.
AC 095N98:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Serum/glucocorticoid-regulated kinase 1 (fragment).
GN SGK1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE-21168084; PubMed-11266509;
RA Mick V.E., Itani O.A., Loftus R.W., Husted R.F., Schmidt T.J.,
RA Thomas C.P.;
RT "The alpha-Subunit of the Epithelial Sodium Channel Is an Aldosterone-
RT Induced Transcript in Mammalian Collecting Ducts, and This
RT Transcriptional Response Is Mediated via Distinct cis-Elements in the
RT 5'-Flanking Region of the Gene.";
RL Mol. Endocrinol. 15:575-588(2001).
DR EMBL: AF317416; AAK54044.1; -.
DR NON_TER
FT 1
SQ SEQUENCE 40 AA; 4189 MW; E54EB18676E225B6 CRC64;

Query Match 84.7%; Score 72; DB 6; Length 40;
Best Local Similarity 87.5%; Pred. No. 2.5e-05;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
:|||||
DB 25 EAFLGFSYAPPMDSXL 40

RESULT 5
093524 PRELIMINARY: PRT; 434 AA.
AC 093524:
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical 49.1 kDa protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE-Renal tubule;
RX MEDLINE-99162637; PubMed-10051674;
RA Chen S.Y., Bhargava A., Mastrobardino L., Meijer O.C., Wang J.,
RA Buse P., Firestone G.L., Verrey F., Pearce D.;
RT "Epithelial sodium channel regulated by aldosterone-induced protein
RT sgk.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 96:2514-2519(1999).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF057138; AAC62398.1; -.
DR HSP: P00517; 1YDR.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase_C; 1.
DR ProDom: PD000001; Prot_Kinase; 1.
DR SMART: SM00220; S_TK_X; 1.
DR SMART: SM00133; S_TK_X; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 434 AA; 49130 MW; 4A061E38BBA6F61 CRC64;

Query Match 82.4%; Score 70; DB 13; Length 434;
Best Local Similarity 75.0%; Pred. No. 0.0074;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
bb 419 EAFMGFSYAPPTDSFL 434
|||||
|||||

RESULT 6
Q8VEK1 PRELIMINARY; PRT; 429 AA.
AC Q8VEK1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to serum/glucocorticoid regulated kinase-like.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: BC018363; AAH18363.1; -.
DR MGD; MGI:2182368; Sgk3.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR001683; PX.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase_C; 1.
DR Pfam: PF00787; PX; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TK_X; 1.
DR SMART: SM00133; S_TK_X; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00195; PX; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 429 AA; 49077 MW; A9C6B7A0C34031F3 CRC64;

Query Match 68.8%; Score 58.5; DB 11; Length 429;
Best Local Similarity 70.6%; Pred. No. 0.074;
Matches 12; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 EAFLGFSYAPPT-DSFL 16
DB 413 DAFVGFYAPPTDSFL 429
|||||
|||||

```

```

RESULT 7
Q97YG2 PRELIMINARY; PRT; 374 AA.
AC Q97YG2;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Transport membrane protein (Permease).
GN SS01361.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyez M.J., Chan-Weiher C.C.-Y., Clausen J.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006750; AAK41596.1; -.
DR InterPro: IPR007114; MFS.
DR PROSITE: PS50850; MFS; 1.
DR Complete proteome.
SQ SEQUENCE 374 AA; 41026 MW; 5B3A8183B71E8F90 CRC64;

Query Match 63.5%; Score 54; DB 17; Length 374;
Best Local Similarity 60.0%; Pred. No. 0.39;
Matches 12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;

QY 1 EAFLGFSYAPPTDSFL 16
DB 290 EAFLGFSYAPPTDSFL 309
|||||
|||||

RESULT 8
Q97X02 PRELIMINARY; PRT; 374 AA.
AC Q97X02;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Transport membrane protein (Permease).
GN SS01958.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyez M.J., Chan-Weiher C.C.-Y., Clausen J.G., Curtis B.A.,
RA De Moors A., Erauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006803; AAK42149.1; -.
DR InterPro: IPR007114; MFS.
DR PROSITE: PS50850; MFS; 1.
DR Complete proteome.
SQ SEQUENCE 374 AA; 41351 MW; 8BB348C7AE34C954 CRC64;

```

Query Match	63.5%; Score 54; DB 17; Length 374;	
Best Local Similarity	60.0%; Pred. No. 0.39;	
Matches	12; Conservative 1; Mismatches 3; Indels 4; Gaps 1;	
QY	1 EAF...GFSYAPPTDSFL 16	
DB	290 EAF...GFSYAPPTDSFL 309	
RESULT 9		
Q972T2		
ID	Q972T2 PRELIMINARY; PRT; 375 AA.	
AC	Q972T2	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE	Hypothetical protein ST1051.	
DE	Hypothetical protein ST1051.	
GN	ST1051.	
OS	Sulfolobus tokodaii.	
OS	Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;	
OC	Sulfolobus.	
OX	NCBI_TaxID=111955;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-JCM 10545 / 7;	
RC	PubMed=11572479;	
RA	Kawarayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,	
RA	Sekine M., Baba S.-I., Ankei A., Kosugi H., Hosoyama A., Fukui S.,	
RA	Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,	
RA	Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki M., Oguchi A.,	
RA	Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,	
RA	Oshima T., Kikuchi H.;	
RT	"Complete genome sequence of an aerobic thermocacidophilic	
RT	Crenarchaeon, Sulfolobus tokodaii strain7."	
RL	DNA Res. 8:123-140(2001).	
DR	EMBL; AP000984; BAB6081.1; -.	
DR	InterPro; IPR007114; MFS.	
DR	PROSITE; PS0850; MFS; 1.	
KN	Hypothetical protein; Complete proteome.	
QY	SEQUENCE 375 AA; 41107 MW; 493D3B5600835125 CRC64;	
Query Match	56.5%; Score 48; DB 17; Length 375;	
Best Local Similarity	55.0%; Pred. No. 4.3;	
Matches	11; Conservative 1; Mismatches 4; Indels 4; Gaps 1;	
QY	1 EAF...GFSYAPPTDSFL 16	
DB	290 EAF...GFSYAPPTDSFL 309	
RESULT 10		
Q9RH02		
ID	Q9RH02 PRELIMINARY; PRT; 164 AA.	
AC	Q9RH02	
DT	01-JUN-2002 (TrEMBLrel. 21, Created)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	Flavodoxin.	
DE	Flavodoxin.	
GN	FN0119.	
OS	Fusobacterium nucleatum (subsp. nucleatum).	
OS	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;	
OC	Fusobacterium.	
OX	NCBI_TaxID=76856;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-ATCC 25586;	
RC	MEDLINE=21886394; PubMed=11899109;	
RA	Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,	
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,	
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,	
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,	
RA	Fonstein M., Kyripides N., Overbeek R.;	

DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
 DE CG12134-PB (RE20857p).
 GN CG12134.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidel-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banjon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreria S., Frise E., Galie R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RA EMBL; AE003831; AAM71048.1;
 DR EMBL; BT001570; AAM71325.1;
 DR FlyBase; FBgn0033471; CG12134.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 3.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PSS0082; WD_REPEATS_2; 1.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Repeat: WD repeat.
 SQ SEQUENCE 341 AA; 37852 MW; CC7A1C4CE9AF0F1 CRC64;
 Query Match 52.98; Score 45; DB 5; Length 341;
 Best Local Similarity 88.9%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 6 FSYAPPTDS 14
 DB 25 FSYLPPTDS 33
 RESULT 13
 Q9V5E5 PRELIMINARY; PRT: 411 AA.
 ID Q9V5E5;
 AC Q9V5E5;
 DT 01-MAY-2000 (TREMELrel. 13, created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
 DE CG12134 protein (RE62785p).
 GN CG12134.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matteli B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.B.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas A., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh K.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarini H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Cealniker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003831; AA558666.1; -;
 DR EMBL; AY071560; AA49182.1; -;
 DR Flybase; FBgn0033471; CG12134.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 3.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PS00082; WD_REPEATS.2; 1.
 DR PROSITE; PS02094; WD_REPEATS_REGION; 1.
 KW Repeat: WD repeat.
 SQ SEQUENCE 411 AA; 45732 MW; FA485EB81498C2F6 CRC64;
 Query Match 52.9%; Score 45; DB 5; Length 411;
 Best Local Similarity 88.9%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 6 FSYAPPTDS 14
 DB 95 FSYLPPTDS 103
 RESULT 14
 Q93YS3 PRELIMINARY; PRT; 896 AA.
 AC Q93YS3
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 99.2 kDa protein.
 GN AT1G69360.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.

RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk K.,
 RA Hayashizaki Y., Ishida J., Jones F., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RT "Full Length cDNA of gene At1g69360 (GI:15222290).";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY059788; AAL24136.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 896 AA; 99237 MW; BEB93E9546FF2D7F CRC64;
 Query Match 52.9%; Score 45; DB 10; Length 896;
 Best Local Similarity 72.7%; Pred. No. 37;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 GFSYAPPTDSF 15
 DB 45 GFSARPTDSY 55
 RESULT 15
 Q9CS20 PRELIMINARY; PRT; 910 AA.
 AC Q9CS20
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical 100.6 kDa protein.
 GN F10D13.5 OR F23010.6.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.E., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.B., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
 RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Nizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";
 RL Nature 408:816-820(2000).
 DR EMBL; AC073178; AAG50100.1; -;
 DR EMBL; AC018364; AAG52500.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 910 AA; 100599 MW; 98941B0D258ED4F4 CRC64;
 Query Match 52.9%; Score 45; DB 10; Length 910;
 Best Local Similarity 72.7%; Pred. No. 37;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 GFSYAPPTDSF 15
 DB 59 GFSARPTDSY 69

Tue Aug 12 10:56:10 2003

us-10-000-039a-4.rspt

Page 7

Search completed: August 11, 2003, 08:22:16
Job time : 6.29185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:14:59 : Search time 2.26609 Seconds
(without alignments)
298.740 Million cell updates/sec

Title: US-10-000-039A-4
Perfect score: 85
Sequence: 1 EAFGLFSYAPPTDSEL 16

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	16	4	US-09-031-295-4
2	85	100.0	431	2	US-08-712-709-5
3	85	100.0	431	3	US-09-111-444-5
4	85	100.0	431	3	US-09-541-228-5
5	85	100.0	431	4	US-09-031-295-2
6	79	92.9	430	2	US-08-712-709-9
7	79	92.9	430	3	US-09-111-444-9
8	79	92.9	430	3	US-09-541-228-9
9	43	50.6	66	4	US-09-430-564-7
10	43	50.6	66	4	US-09-430-564-13
11	43	50.6	482	4	US-08-430-564-2
12	43	50.6	495	2	US-08-749-902-3
13	43	50.6	495	4	US-09-430-564-3
14	41	48.2	162	1	US-08-624-125-5
15	41	48.2	162	1	US-08-624-125-8
16	41	48.2	162	2	US-08-663-310-4
17	41	48.2	162	2	US-08-663-310-11
18	41	48.2	162	2	US-09-006-491-4
19	41	48.2	162	2	US-09-006-491-11
20	41	48.2	162	3	US-09-335-919-4
21	41	48.2	162	3	US-09-335-919-11
22	41	48.2	162	3	US-08-980-832-34
23	41	48.2	162	4	US-08-937-155-5
24	41	48.2	162	4	US-08-937-155-8
25	40	47.1	445	4	US-09-252-991A-28586
26	40	47.1	724	4	US-09-562-737-23
27	39	45.9	525	2	US-08-749-902-7

28	39	45.9	525	2	US-08-749-902-8	Sequence 8, Appli
29	39	45.9	525	4	US-09-430-564-16	Sequence 16, Appli
30	39	45.9	830	4	US-09-562-737-32	Sequence 32, Appli
31	39	45.9	831	1	US-08-073-384C-5	Sequence 5, Appli
32	39	45.9	831	1	US-08-254-359A-5	Sequence 5, Appli
33	39	45.9	831	1	US-08-483-043-5	Sequence 5, Appli
34	39	45.9	831	1	US-08-481-238-5	Sequence 5, Appli
35	39	45.9	831	2	US-08-471-066B-5	Sequence 5, Appli
36	39	45.9	831	2	US-08-484-956-5	Sequence 5, Appli
37	39	45.9	831	2	US-08-757-653-5	Sequence 5, Appli
38	39	45.9	831	2	US-08-599-491-5	Sequence 5, Appli
39	39	45.9	831	2	US-08-756-386-5	Sequence 5, Appli
40	39	45.9	831	2	US-08-823-516-5	Sequence 5, Appli
41	39	45.9	831	3	US-08-682-853A-5	Sequence 5, Appli
42	39	45.9	831	3	US-08-759-038-5	Sequence 5, Appli
43	39	45.9	831	3	US-08-758-314-5	Sequence 5, Appli
44	39	45.9	831	4	US-09-350-309-5	Sequence 5, Appli
45	39	45.9	831	4	US-08-520-946-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-031-295-4
: Sequence 4, Application US/09031295
: Patent No. 6326181

: GENERAL INFORMATION:
: APPLICANT: LANG, Florian
: APPLICANT: WALDEGGER, Tubingen
: TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SUK
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY & LARDNER
: STREET: 3000 K Street, N.W.
: CITY: Washington
: STATE: D.C.

: COUNTRY: U.S.A.
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/031.295
: FILING DATE: 26-FEB-1998

: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: DE 197-08-173.8
: FILING DATE: 28-FEB-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Sandercock, Colin G.
: REGISTRATION NUMBER: 31,298
: REFERENCE/DOCKET NUMBER: 058315/0123

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 672-5300
: TELEFAX: (202) 672-5399
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 16 amino acids
: TYPE: amino acid
: STRANDEDNESS:

: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-031-295-4

Query Match 100.0% Score 85; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EAFGLFSYAPPTDSEL 16
|||||

Db 1 EAFLGFSYAPPTDSFL 16

RESULT 2

US-08-712-709-5

; Sequence 5, Application US/08712709

; Patent No. 5863780

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Guegler, Karl J.

; APPLICANT: Hawkins, Phillip R.

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/712,709

; FILING DATE: Filed Herewith

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0118 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 431 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; LIBRARY: Consensus

; CLONE: Consensus

; US-08-712-709-5

Query Match 100.0%; Score 85; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 8.8e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16

Db 416 EAFLGFSYAPPTDSFL 431

RESULT 3

US-09-111-444-5

; Sequence 5, Application US/09111444

; Patent No. 6045792

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Guegler, Karl J.

; APPLICANT: Hawkins, Phillip R.

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/111,444

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/712,709

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0118 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 431 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; LIBRARY:

; CLONE: Consensus

; US-09-111-444-5

Query Match 100.0%; Score 85; DB 3; Length 431;

Best Local Similarity 100.0%; Pred. No. 8.8e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16

Db 416 EAFLGFSYAPPTDSFL 431

RESULT 4

US-09-541-228-5

; Sequence 5, Application US/09541228

; Patent No. 6232077

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Guegler, Karl J.

; APPLICANT: Hawkins, Phillip R.

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/541,228

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/712,709

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0118 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 431 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY:
 CLONE: Consensus
 US-09-541-228-5

Query Match 100.0%; Score 85; DB 3; Length 431;
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 416 EAFLGFSYAPPTDSFL 431

RESULT 5
 US-09-031-295-2
 Sequence 2, Application US/09031295
 Patent No. 6326181

GENERAL INFORMATION:
 APPLICANT: LANG, Florian
 APPLICANT: WALDEGER, Tubingen
 TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY & LARDNER
 STREET: 3000 K Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/031,295
 FILING DATE: 26-FEB-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 197-08-173.8
 FILING DATE: 28-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Sandercock, Colin G.
 REGISTRATION NUMBER: 31,298
 REFERENCE/DOCKET NUMBER: 058315/0123
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300
 TELEFAX: (202) 672-5399
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 431 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-031-295-2

Query Match 100.0%; Score 85; DB 4; Length 431;
 Best Local Similarity 100.0%; Pred. No. 8.8e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 416 EAFLGFSYAPPTDSFL 431

RESULT 6
 US-08-712-709-9
 Sequence 9, Application US/08712709
 Patent No. 5863780
 GENERAL INFORMATION:
 APPLICANT: Au-Young, Janice
 APPLICANT: Guegler, Karl J.
 APPLICANT: Hawkins, Phillip R.
 TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/712,709
 FILING DATE: Filed Herewith
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0118 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 430 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 294637
 US-08-712-709-9

Query Match 92.9%; Score 79; DB 2; Length 430;
 Best Local Similarity 93.8%; Pred. No. 7.6e-05;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EAFLGFSYAPPTDSFL 16
 |||||
 DB 415 EAFLGFSYAPPTDSFL 430

RESULT 7
 US-09-111-444-9
 Sequence 9, Application US/09111444
 Patent No. 6045792
 GENERAL INFORMATION:
 APPLICANT: Au-Young, Janice
 APPLICANT: Guegler, Karl J.
 APPLICANT: Hawkins, Phillip R.
 TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,444
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637
US-09-111-444-9

Query Match 92.9%; Score 79; DB 3; Length 430;
Best Local Similarity 93.8%; Pred. No. 7.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EAFLGFSYAPPTDSFL 16
|||||
Db 415 EAFLGFSYAPPMDSFL 430

RESULT 8
US-09-341-228-9
Sequence 9, Application US/09541228
Patent No. 6232077
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637
US-09-541-228-9

Query Match 92.9%; Score 79; DB 3; Length 430;
Best Local Similarity 93.8%; Pred. No. 7.6e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EAFLGFSYAPPTDSFL 16
|||||
Db 415 EAFLGFSYAPPMDSFL 430

RESULT 9
US-09-430-564-7
Sequence 7, Application US/09430564
Patent No. 6372467
GENERAL INFORMATION:
APPLICANT: John Blenis
APPLICANT: Kay K. Lee-Fruman
APPLICANT: Calvin J. Kuo
TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
PRIMERS, PROBES, AND DETECTION METHODS
FILE REFERENCE: 00246/506002
CURRENT APPLICATION NUMBER: US/09/430,564
CURRENT FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/106,141
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 66
TYPE: PRT
ORGANISM: Homo sapiens
US-09-430-564-7

Query Match 50.6%; Score 43; DB 4; Length 66;
Best Local Similarity 58.3%; Pred. No. 5.2;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EAFLGFSYAPPT 12
|||||
Db 51 QAFLGFTYVAPS 62

RESULT 10
US-09-430-564-13
Sequence 13, Application US/09430564
Patent No. 6372467
GENERAL INFORMATION:
APPLICANT: John Blenis
APPLICANT: Kay K. Lee-Fruman
APPLICANT: Calvin J. Kuo
TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
PRIMERS, PROBES, AND DETECTION METHODS
FILE REFERENCE: 00246/506002
CURRENT APPLICATION NUMBER: US/09/430,564
CURRENT FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/106,141
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 66
TYPE: PRT
ORGANISM: Homo sapiens

US-09-430-564-13

Query Match 50.6%; Score 43; DB 4; Length 66;
 Best Local Similarity 58.3%; Pred. No. 5.2;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EAFGLFSYAPPT 12
 :|||||: 1;
 Db 51 QAFGLFTYVAPS 62

RESULT 11

US-09-430-564-2
 ; Sequence 2, Application US/09430564
 ; Patent No. 6372467
 ; GENERAL INFORMATION:
 ; APPLICANT: John Blenis
 ; APPLICANT: Kay K. Lee-Fruman
 ; APPLICANT: Calvin J. Kuo
 ; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
 ; FILE REFERENCE: 00246/506002
 ; CURRENT APPLICATION NUMBER: US/09/430,564
 ; CURRENT FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 60/106,141
 ; PRIOR FILING DATE: 1998-10-29
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 482
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-430-564-2

Query Match 50.6%; Score 43; DB 4; Length 482;
 Best Local Similarity 58.3%; Pred. No. 38;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EAFGLFSYAPPT 12
 :|||||: 1;
 Db 382 QAFGLFTYVAPS 393

RESULT 12

US-08-749-902-3
 ; Sequence 3, Application US/08749902
 ; Patent No. 5985635
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Hillman, Jennifer L.
 ; TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/749,902
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0150 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 495 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: Consensus
 US-08-749-902-3

Query Match 50.6%; Score 43; DB 2; Length 495;
 Best Local Similarity 58.3%; Pred. No. 39;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EAFGLFSYAPPT 12
 :|||||: 1;
 Db 385 QAFGLFTYVAPS 396

RESULT 13

US-09-430-564-3
 ; Sequence 3, Application US/09430564
 ; Patent No. 6372467
 ; GENERAL INFORMATION:
 ; APPLICANT: John Blenis
 ; APPLICANT: Kay K. Lee-Fruman
 ; APPLICANT: Calvin J. Kuo
 ; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
 ; FILE REFERENCE: 00246/506002
 ; CURRENT APPLICATION NUMBER: US/09/430,564
 ; CURRENT FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 60/106,141
 ; PRIOR FILING DATE: 1998-10-29
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 495
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-430-564-3

Query Match 50.6%; Score 43; DB 4; Length 495;
 Best Local Similarity 58.3%; Pred. No. 39;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EAFGLFSYAPPT 12
 :|||||: 1;
 Db 395 QAFGLFTYVAPS 406

RESULT 14

US-08-624-125-5
 ; Sequence 5, Application US/08624125
 ; Patent No. 5744341
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; APPLICANT: SUN ZAIREN
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; ADDRESSEE: P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
 ; CITY: ARLINGTON
 ; STATE: VA

;; COUNTRY: USA
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/624,125
;; FILING DATE: 29-MAR-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KELBER, STEVEN B.
;; REGISTRATION NUMBER: 30,073
;; REFERENCE/DOCKET NUMBER: 2747-063-27
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 162 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-624-125-5

Query Match 48.2%; Score 41; DB 1; Length 162;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 GFSYAPPTD 13
Db 134 GFIYAPPVD 142

RESULT 15
US-08-624-125-8
; Sequence 8, Application US/08624125
; Patent No. 5744341
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/624,125
;; FILING DATE: 29-MAR-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KELBER, STEVEN B.
;; REGISTRATION NUMBER: 30,073
;; REFERENCE/DOCKET NUMBER: 2747-063-27
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 162 amino acids

;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-624-125-8

Query Match 48.2%; Score 41; DB 1; Length 162;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 GFSYAPPTD 13
Db 134 GFIYAPPVD 142

Search completed: August 11, 2003, 08:24:45
Job time : 3.26609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:24:50 ; Search time 448 Seconds
(without alignments)
1461.058 Million cell updates/sec

Title: US-10-000-039a-4
Perfect score: 85
Sequence: 1 EAFGLFSYAPPTDSFL 16

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n model -DEV=slp
-Q/cgn21/USPTO.spool/US1000039/runat_01082003_085418_27511/app_query.fasta_1.661
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.coi -LIST=45
-DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPT=US1000039.cgn_1.1.5569.0/runat_01082003_085418_27511 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb.ba: *
2: gb.htg: *
3: gb.in: *
4: gb.om: *
5: gb.ov: *
6: gb.pat: *
7: gb.ph: *
8: gb.pl: *
9: gb.pr: *
10: gb.ro: *
11: gb.sts: *
12: gb.sy: *
13: gb.un: *
14: gb.vi: *
15: em.ba: *
16: em.fun: *
17: em.hum: *
18: em.in: *
19: em.mu: *
20: em.om: *
21: em.or: *
22: em.ov: *
23: em.pat: *
24: em.ph: *
25: em.pl: *
26: em.ro: *
27: em.sts: *
28: em.un: *

29: em.vi: *
30: em.htg.hum: *
31: em.htg.in: *
32: em.htg.other: *
33: em.htg.mus: *
34: em.htg.pln: *
35: em.htg.ro: *
36: em.htg.mam: *
37: em.htg.vrt: *
38: em.sy: *
39: em.htg.hum: *
40: em.htg.mus: *
41: em.htg.other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	694	6	AX201922 Sequence
2	85	100.0	1296	6	AX056374 Sequence
3	85	100.0	2281	6	AX017284 Sequence
4	85	100.0	2281	6	AX524988 Sequence
5	85	100.0	2281	6	BD134455 Human nuc
6	85	100.0	2311	6	AR151390 Sequence
7	85	100.0	2311	6	AR270209 Sequence
8	85	100.0	2311	6	BD021928 Human pro
9	85	100.0	2354	6	AX553549 Sequence
10	85	100.0	2364	6	AK098509 Homo sapi
11	85	100.0	2370	6	AR179441 Sequence
12	85	100.0	2370	6	AX002570 Sequence
13	85	100.0	2370	6	AX337834 Sequence
14	85	100.0	2370	6	AX411211 Sequence
15	85	100.0	2372	9	HSRNASRTPK
16	85	100.0	2380	9	AF153609 Homo sapi
17	85	100.0	2382	9	BC001263 Homo sapi
18	85	100.0	3196	6	AX713538 Sequence
19	85	100.0	3196	9	AK055077 Homo sapi
20	85	100.0	5718	9	HSSGK
21	85	100.0	113673	9	AL135839 Human DNA
22	80	94.1	431	6	AX305574 Sequence
23	80	94.1	1659	10	BC002222 Mus muscu
24	80	94.1	2423	10	BC005720 Mus muscu
25	80	94.1	2426	10	AF205855 Mus muscu
26	80	94.1	2429	10	AF139638 Mus muscu
27	80	94.1	2432	6	AX056375 Sequence
28	80	94.1	208405	2	AC114405 Mus muscu
29	79	92.9	1920	4	AF139639 Oryctolag
30	79	92.9	2435	10	RATSGPK
31	79	92.9	142609	2	AC118095 Rattus no
32	79	92.9	220350	2	AC107530 Rattus no
33	79	92.9	240716	2	AC112280 Rattus no
34	76	89.4	2470	5	SAC223715 Squalus a
35	76	89.4	3105	5	SAC223716 Squalus a
36	72	84.7	151	4	AF317416 Canis fam
37	70	82.4	1417	5	AF057138 Xenopus l
38	68	80.0	2339	5	BC052134 Danio rer
39	68	80.0	180682	2	BC323831 Danio rer
40	67.5	79.4	225264	2	BC470095 Danio rer
41	64	75.3	183140	10	AC124486 Mus muscu
42	58.5	68.8	606	6	E32370 Serine-thr
43	58.5	68.8	1768	9	BC015326 Homo sapi
44	58.5	68.8	1803	10	AF312007 Mus muscu
45	58.5	68.8	2250	6	AX056377 Sequence

ALIGNMENTS

RESULT 1

AX201922
 LOCUS AX201922 694 bp DNA linear PAT 30-AUG-2001
 DEFINITION Sequence 54 from Patent WO0153524.
 ACCESSION AX201922
 VERSION AX201922.1 GI:15391773
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Rees, R.C., Li, G. and Mian, S.
 TITLE Cancer associated genes and their products
 JOURNAL Patent: WO 0153524-A 54 26-JUL-2001;
 The Nottingham Trent University (GB)
 FEATURES
 source
 1..694
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 148 a 153 c 172 g 203 t 18 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.49e-06 Length: 694
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-000-039a-4 (1-16) x AX201922 (1-694)
 Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 112 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGAGCTCTTTCTC 159
 RESULT 2
 AX056374
 LOCUS AX056374 1296 bp DNA linear PAT 13-JAN-2001
 DEFINITION Sequence 18 from Patent WO0073469.
 ACCESSION AX056374
 VERSION AX056374.1 GI:12229081
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.
 TITLE Protein kinases
 JOURNAL Patent: WO 0073469-A 18 07-DEC-2000;
 Sugen, Inc. (US)
 FEATURES
 source
 1..1296
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 345 a 333 c 293 g 325 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.23e-05 Length: 1296
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-000-039a-4 (1-16) x AX056374 (1-1296)
 Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1246 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGAGCTCTTTCTC 1293
 RESULT 3
 AX017284
 LOCUS AX017284 2281 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 39 from Patent WO9947669.
 ACCESSION AX017284
 VERSION AX017284.1 GI:10042202
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Schmitt, A., Specht, T., Dahl, E., Hinzmann, B., Rosenthal, A. and
 Pilarsky, C.
 TITLE Human nucleic acid sequences from tissue of breast tumors
 JOURNAL Patent: WO 9947669-A 39 23-SEP-1999;
 SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
 BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
 (DE); PILARSKY CHRISTIAN (DE)
 FEATURES
 source
 1..2281
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 601 a 498 c 494 g 688 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.18e-05 Length: 2281
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-10-000-039a-4 (1-16) x AX017284 (1-2281)
 Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1188 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGAGCTCTTTCTC 1245
 RESULT 4
 AX524988
 LOCUS AX524988 2281 bp DNA linear PAT 21-NOV-2002
 DEFINITION Sequence 39 from Patent EP1236799.
 ACCESSION AX524988
 VERSION AX524988.1 GI:25170070
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Specht, T., Hinzmann, B., Schmitt, A., Pilarsky, C., Dahl, E. and
 Rosenthal, A.
 TITLE Human nucleic acid sequences derived from breast tumor tissue
 JOURNAL Patent: EP 1236799-A 39 04-SEP-2002;
 metaGen Pharmaceuticals GmbH (DE)
 FEATURES
 source
 1..2281
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 601 a 498 c 494 g 688 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.18e-05 Length: 2281
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-000-039a-4 (1-16) x AX524988 (1-2281)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 |||||||
 Db 1188 GAGGCTTTCCTAGGCTTTTCTATGCGCTCCACGAGACTCTTTCTC 1235

RESULT 5
 BD134455 2281 bp DNA linear PAT 18-SEP-2002
 LOCUS Human nucleic acid sequence originating in mammary tumor tissue.
 DEFINITION
 ACCESSION BD134455
 VERSION BD134455.1 GI:23229400
 KEYWORDS JP 2002506643-A/33.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Sperm, T., Hintzman, B., Armin, S., Pirarski, C., Edgar, D. and Rosenthal, A.
 Human nucleic acid sequence originating in mammary tumor tissue
 Patent: JP 2002506643-A 33 05-MAR-2002;
 METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
 OS Homo sapiens (human)
 PN JP 2002506643-A/33
 PD 05-MAR-2002
 PF 19-MAR-1999 JP 2000536852
 PR 20-MAR-1998 DE 198 13 839.3
 PI THOMAS SPEFT, BERND HINTZMAN, SCHOEITZ ARMIN, CHRISTIAN PIRARSKI, DUHL EDGAR.
 PI ANDRE ROSENTHAL
 PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P35/00, C07K14/47,
 PC C07K16/18, C12N15/10, C12N15/00, A61K37/02, C12N5/00 CC Human
 nucleic acid sequence originating in mammary tumor CC
 tissue
 FH key Location/Qualifiers
 FT source 1..2281 /organism="Homo sapiens (human)"
 FEATURES
 source
 Location/Qualifiers
 1..2281 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 601 a 498 c 494 g 688 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,18e-05 Length: 2281
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-000-039a-4 (1-16) x BD134455 (1-2281)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 |||||||
 Db 1188 GAGGCTTTCCTAGGCTTTTCTATGCGCTCCACGAGACTCTTTCTC 1235

RESULT 6
 AR151390 2311 bp DNA linear PAT 08-AUG-2001
 LOCUS Sequence 6 from patent US 6232077.
 DEFINITION
 ACCESSION AR151390
 VERSION AR151390.1 GI:15117440
 KEYWORDS
 SOURCE unknown.
 ORGANISM Unknown.

Unclassified.
 1 (bases 1 to 2311)
 Au-Young, J., Guegler, K.J. and Hawkins, P.R.
 Human protein kinases
 Patent: US 6232077-A 6 15-MAY-2001;
 Location/Qualifiers
 1..2311 /organism="unknown"
 BASE COUNT 604 a 508 c 506 g 692 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,21e-05 Length: 2311
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-000-039a-4 (1-16) x AR151390 (1-2311)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 |||||||
 Db 1255 GAGGCTTTCCTAGGCTTTTCTATGCGCTCCACGAGACTCTTTCTC 1302

RESULT 7
 AR270209 2311 bp DNA linear PAT 10-APR-2003
 LOCUS Sequence 772 from patent US 6500938.
 DEFINITION
 ACCESSION AR270209
 VERSION AR270209.1 GI:29701443
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 1 (bases 1 to 2311)
 Au-Young, J. and Seilhamer, J.J.
 Composition for the detection of signaling pathway gene expression
 Patent: US 6500938-A 772 31-DEC-2002;
 Location/Qualifiers
 1..2311 /organism="unknown"
 BASE COUNT 604 a 508 c 506 g 692 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,21e-05 Length: 2311
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-000-039a-4 (1-16) x AR270209 (1-2311)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 |||||||
 Db 1255 GAGGCTTTCCTAGGCTTTTCTATGCGCTCCACGAGACTCTTTCTC 1302

RESULT 8
 BD021928 2311 bp DNA linear PAT 27-AUG-2002
 LOCUS Human protein kinases.
 DEFINITION
 ACCESSION BD021928
 VERSION BD021928.1 GI:22563151
 KEYWORDS JP 2001506843-A/3.
 SOURCE unidentified
 ORGANISM unidentified.
 Unclassified.
 1 (bases 1 to 2311)
 Young, J.O., Guegler, K.J. and Hopkins, P.R.
 Human protein kinases
 Patent: JP 2001506843-A 3 29-MAY-2001;
 Location/Qualifiers
 1..2311 /organism="unknown"

```

INSITE PHARMACEUTICALS INC
PN JP 2001508643-A/3
PD 29-MAY-2001
PF 12-SEP-1997 JP 1998513776
PR 12-SEP-1996 US 08/712709
PI JANICE O YOUNG, KARL J GEGURA, PHILLIP R HORKINS PC
C12N15/00, A61K38/45, A61K39/395, A61K45/00, A61K48/00, PC
A61P11/06,
PC A61P19/02, A61P25/28, A61P29/00, A61P35/00, A61P43/00, C07K15/40,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/12, C12Q1/48, C12N15/00, C12N5/
CC 00, A61K37/52
FH Key Location/Qualifiers.
FEATURES source
1..2311 Location/Qualifiers
/mol_type="unidentified"
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 604 a 507 c 507 g 692 t 1 others
ORIGIN
Alignment Scores: 2.21e-05 Length: 2311
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-4 (1-16) x BD021928 (1-2311)
Qy 1 GluAlaPheLeuGlyPheSerTyAlaProProThrasPheLeu 16
1255 GAGGCTTTCTAGGCTTTCTATGCGCCCTCCACGAGCTTTTCCTC 1302
RESULT 9
AX553549
LOCUS AX553549 2354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 1 from Patent WO02074987.
ACCESSION AX553549
VERSION AX553549.1 GI:25897549
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lang, F., Busjahn, A. and Luft, F.C.
Quantitative diagnostic analysis of hypertonia
Patent: WO 02074987-A 1 26-SEP-2002;
Lang, Florian (DE)
FEATURES source
1..2354 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
43..1338
/feature="unnamed protein product"
/codon_start=1
/db_xref="GI:25897550"
/translaton="MTVKTEAAAGTLTYSMRGVAILTAPMKQRKRGKLNDFQITAN
NSYACKPEQVSIKITSQPEPELMNANPPSPSQQLNGSSPHAKPSDFHFLK
VIGSGFQVLLARHAEVYAVKVLQKAILKKEKHIMSERVNLKNYKHFLV
GLHFSQADKLYFVLDYINGGELFVHLQERCFLPEPRAFVAAEIASALGVILHSLN
VYRLKPGVILLDSQHVILDFGLCKENTENHSTTSTFCGTPPEYLAPEVLHKKPYDR
TVDWCKLGVILLYGLPFFYSRNTAEVNTLNKFLQIKPNTITNSARHLLGLQK
DRTRELKAKDDFVEIKSHVESFLINWDDLINXKITPPFPNPGSGNOLRHFPEFTEE
PVPNSIGKSDSVLVTSYKEREAEFLGFSIAPPTDSFL"
762
variation
/Note="1. SNP (C in T), stumme Mutation, d.h. beide
Versionen des SNPs resultieren in der AMINO ACID Asp in

```

```

BASE COUNT 613 a 518 c 513 g 710 t
ORIGIN

```

```

Alignment Scores: 2.25e-05 Length: 2354
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-4 (1-16) x AX553549 (1-2354)
Qy 1 GluAlaPheLeuGlyPheSerTyAlaProProThrasPheLeu 16
1288 GAGGCTTTCTAGGCTTTCTATGCGCCCTCCACGAGCTTTTCCTC 1335

```

```

RESULT 10
AK098509
LOCUS AK098509 2364 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ25643 fis, clone STM07148, highly similar to
SERINE/THREONINE-PROTEIN KINASE SGK (EC 2.7.1.-).
ACCESSION AK098509
VERSION AK098509.1 GI:21758535
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1
AUTHORS Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
Ono, Y., Hotuta, T., Hiraoka, S., Muraoka, K., Takiguchi, S.,
Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S.,
Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Ise, T. and
Sugano, S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2364)
Sugano, S. and Suzuki, Y.
Direct Submission
Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)

```

```

TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2364)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
Fax: 81-3-5449-5416)

```

```

COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.

```

```

FEATURES source
1..2364 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM07148"
/tissue="stomach mucosa"
/clone_lib="STM"
/Note="cloning vector: pME18SFL3"
BASE COUNT 615 a 524 c 517 g 708 t
ORIGIN

```

```

Alignment Scores: 2.26e-05 Length: 2364
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```



```

DB:
US-10-000-039A-4 (1-16) x AK098509 (1-2364)
QY 1 GluAlaPheLeuGlyPheSerTyraIaProProThrAspSerPheLeu 16
DB 1303 GAGGCTTTCCTAGGCTTTTCCTATGCGCTCCCGAGGACTCTTCTC 1335

RESULT 11
LOCUS ARI79441 2370 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6326181.
ACCESSION ARI79441
VERSION ARI79441.1 GI:20220996
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2370)
AUTHORS Lang,F. and Waldegger,S.
TITLE Cell volume-regulated human kinase h-sgk
JOURNAL Patent: US 6326181-A 1 04-DEC-2001.
FEATURES
    source
        location/Qualifiers
            1..2370
                /organism="unknown"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 2.27e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-000-039A-4 (1-16) x ARI79441 (1-2370)
QY 1 GluAlaPheLeuGlyPheSerTyraIaProProThrAspSerPheLeu 16
DB 1288 GAGGCTTTCCTAGGCTTTTCCTATGCGCTCCCGAGGACTCTTCTC 1335

RESULT 12
LOCUS AX002570 2370 bp DNA linear PAT 10-MAR-2000
DEFINITION Sequence 1 from Patent EP0861896.
ACCESSION AX002570
VERSION AX002570.1 GI:7242111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2370)
AUTHORS Lang,F.P. and Waldegger,S.D.
JOURNAL Cell volume regulated human kinase h-sgk
PATENT: EP 0861896-A 1 02-SEP-1998;
DADE BEHRING MARBURG GMBH (DE)
FEATURES
    source
        location/Qualifiers
            1..2370
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /strain="CAUCASIAN"
                /db_xref="taxon:9606"
                /clone="HSRNAS1PK"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 2.27e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
DB: Gaps: 0

US-10-000-039A-4 (1-16) x AX002570 (1-2370)
QY 1 GluAlaPheLeuGlyPheSerTyraIaProProThrAspSerPheLeu 16
DB 1288 GAGGCTTTCCTAGGCTTTTCCTATGCGCTCCCGAGGACTCTTCTC 1335

RESULT 13
LOCUS AX337834 2370 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 8343 from Patent WO0194629.
ACCESSION AX337834
VERSION AX337834.1 GI:18128553
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2370)
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
JOURNAL Horrigan,S., Soppet,D.R. and Weaver,Z.
FEATURES Cancer gene determination and therapeutic screening using signature
    source
        location/Qualifiers
            1..2370
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 2.27e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-000-039A-4 (1-16) x AX337834 (1-2370)
QY 1 GluAlaPheLeuGlyPheSerTyraIaProProThrAspSerPheLeu 16
DB 1288 GAGGCTTTCCTAGGCTTTTCCTATGCGCTCCCGAGGACTCTTCTC 1335

RESULT 14
LOCUS AX411211 2370 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3858 from Patent WO0229103.
ACCESSION AX411211
VERSION AX411211.1 GI:21443916
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2370)
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
JOURNAL Gene expression profiles in liver cancer
PATENT: WO 0229103-A 3858 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
    source
        location/Qualifiers
            1..2370
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /note="EMBL/GenBank Accession No. Y10032"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN

```

Alignment Scores:

Pred. No.: 2,27e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-4 (1-16) x AX411211 (1-2370)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
|||||
Db 1288 GAGGCTTTCTAGGCTTTCTATAGCGCTTCCACGAGACTCTTTCTCTC 1335

RESULT 15

HSRNASTPK 2370 bp mRNA linear PRI 16-MAY-1997
LOCUS HSRNASTPK
DEFINITION H. sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION Y10032
VERSION Y10032.1 GI:1834510
KEYWORDS serine/threonine protein kinase; sgk gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Waldegger, S., Barth, P., Raber, G. and Lang, F.
TITLE Cloning and characterization of a putative human serine/threonine
protein kinase transcriptionally modified during anisotonic and
isotonic alterations of cell volume
PROC. Natl. Acad. Sci. U.S.A. 94 (9), 4440-4445 (1997)

JOURNAL

MEDLINE 97272242

PUBMED 9114008

REFERENCE 2 (bases 1 to 2370)

AUTHORS Waldegger, S.

TITLE Direct Submission

JOURNAL Submitted (11-DEC-1996) S. Waldegger, University of Tuebingen,

Physiology I, Gmelinstr. 5, D-72076 Tuebingen, FRG

FEATURES

source

1..2370

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="adult"

1..2370

/gene="sgk"

43..1338

/gene="sgk"

/codon_start=1

/product="serine/threonine protein kinase"

/protein_id="CAA71138.1"

/db_xref="GI:1834511"

/db_xref="SWISS-PROT:O00141"

/translation="MTVKEAAGTLTYSRMGMVAIIAFMKQRMGLNDFIQKIAN

NSVAKHPEVQSILKISQPEPELMNANSPSPSQINLGPSSNPHAKPSDFHFLK

VIGKSGFKVLLARHAEVYAVKVLQKAIKKEKHIMSERNVLLKNVKHPLV

GLHFSFOADKLYFVLDYINGGELFYHLOERCQFLERARFYAAEIASALGYLSLNI

TVRDLKPEHILLDSOGHIVLTDFGLCKENIEHNSSTSTFCGTPEYLAPEVLHKQPYDR

TYDWKLCGLAVLYEMLYGLPPYSRNTAEMTDNILNRPQLKPNITNSARHLLEGLOK

DRTKRLGAKDDFMEIKSHVFFSLINDDLLINKITPPFPNVSGPNELRHDFPEFTEE

PVPNSIGKSPDSVLVTASVKAEAEFLGSYPAPPTDSFL"

BASE COUNT 636 a 517 c 513 g 704 t

Alignment Scores:

Pred. No.: 2,27e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-000-039A-4 (1-16) x HSRNASTPK (1-2370)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
|||||
Db 1288 GAGGCTTTCTAGGCTTTCTATAGCGCTTCCACGAGACTCTTTCTCTC 1335

Search completed: August 11, 2003, 10:05:07
Job time : 451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:23:34 ; Search time 59,6364 Seconds
(without alignments)
724.239 Million cell updates/sec

Title: US-10-000-039A-4
Perfect score: 85
Sequence: 1 EARLGFSVAPPTDSFL 16

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q/cn2_1/USF10_spo01/US10000039/runat_01082003_085418_27503/app_query.fasta_1.661
-DB-N_Geneseq_19Jun03 -QFMT-fastap -SUFFIX-rng -MINMATCH-0.1 -LOOPEL-0
-LOOPEXT-0 -UNITS-Bits -START-1 -END-1 -MATRIX-BLOSUM62 -FRANS-human40.cdi
-LIST-45 -DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US10000039 @CGN_1.1.740 @runat_01082003_085418_27503 -NCPUP-6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	85	100.0	694	22	AAH42835 Nucleotide sequenc
2	85	100.0	1296	22	AAF44639 Novel protein kina
3	85	100.0	2281	20	AAZ33647 Human breast tumo
4	85	100.0	2311	19	AAV23833 Human protein kina
5	85	100.0	2311	25	ACA56174 Human signalling p
6	85	100.0	2343	25	ABX74395 Human cDNA sequenc
7	85	100.0	2346	20	AAV74190 Human cell-volume
8	85	100.0	2370	19	AAV48311 Human cDNA differe
9	85	100.0	2370	24	ABN97360 Gene #3858 used to
10	85	100.0	2370	24	ABN97360 Pancreas cancer re
11	85	100.0	2370	24	ABL70006 Mouse ischaemic co
12	80	94.1	431	24	ABF199409 Novel protein kina
13	80	94.1	2432	22	AAF44640 Bovine EST associa
14	79	92.9	404	25	ABX42261 DNA encoding a rat
15	79	92.9	2365	21	AA446592 Bovine PST associa
16	66	77.6	433	25	ABX38963 Human pancreatic c
17	58.5	68.8	615	24	ABV96134 cDNA encoding nove
18	58.5	68.8	851	22	AA527021 DNA encoding novel
19	58.5	68.8	851	23	ABK43988 Novel protein kina
20	58.5	68.8	2250	22	AAF44642 Human serum and ql
21	58.5	68.8	2404	21	AAZ27858 Human full-length
22	58.5	68.8	2483	22	AAK94684 Human serum glucoc
23	58.5	68.8	2512	20	AAV99653 DNA encoding novel
24	58.5	68.8	2572	23	ABK43712 Human polynucleoti
25	58.5	68.8	2702	22	AAI59776 Human serine threo
26	58.5	68.8	2711	22	AAI57990 Human spliced tran
27	58.5	68.8	3019	21	AAAG9248 Human protein enco
28	57	67.1	60	24	ABN35484 Novel protein kina
29	56	65.9	1333	22	AAH99535 cDNA encoding nove
30	56	65.9	1346	22	AA444641 Novel protein kina
31	56	65.9	1366	22	AA544987 Human serum and ql
32	56	65.9	1812	22	RAF44737 Human serum and ql
33	56	65.9	1812	24	AAZ36141 Human serum and ql
34	56	65.9	1834	21	AAZ27856 Human colon specif
35	56	65.9	2146	21	AAZ27856 Human genome fragm
36	56	65.9	2146	21	AAZ27856 Human cDNA clone (
37	52	61.2	3872	24	ABA91312 Human breast cell
38	48	56.5	218	15	AAQ76626 Human foetal liver
39	47	55.3	643	22	AAH03224 Probe #2430 for ge
C 40	46	54.1	478	22	ABA54215 Human brain expres
C 41	46	54.1	478	22	ABA54215 Human bone marrow
C 42	46	54.1	478	22	ABA23964 Probe #2442 for ge
C 43	46	54.1	478	22	AAK02493
C 44	46	54.1	478	22	AAK27931
C 45	46	54.1	478	22	AAI12509

ALIGNMENTS

RESULT 1
AAH42835
ID AAH42835 standard; DNA: 694 BP.
XX
AC AAH42835;
XX
DT 01-OCT-2001 (first entry)
XX
DE Nucleotide sequence of a human sgk protein kinase.
XX
KW Cancer associated gene; prostate cancer; stomach cancer;
KW oesophageal cancer; cancer; vaccine; ss.
XX
OS Homo sapiens.
XX
PN WO200153524-A2.
XX

Tue Aug 12 10:56:07 2003

us-10-000-039a-4.rng

PD 26-JUL-2001.
XX
PF 18-JAN-2001: 2001WO-GB00188.
XX
XX 18-JAN-2000: 2000GB-0000993.
PR
XX (UYN0-) UNIV NOTTINGHAM TRENT.
PA
XX
PI Rees RC, Li G, Mian S;
XX
XX WPI: 2001-476121/51.
DR
XX
PT Use of cancer-associated genes and their products in detecting,
XX monitoring, treating or preventing cancer, specifically prostate
PT cancer, and in developing DNA-based vaccines that promote anti-tumor
PT responses
XX
XX Claim 1: Page 29; 43pp; English.
PS
XX AAH42781-AAH42846 represent the nucleotide sequences of cancer
XX associated genes, identified using SEREX (Serological identification of
XX antigens by Recombinant Expression Cloning). The genes are overexpressed
XX in prostate cancer, and some are overexpressed in other cancers such as
XX stomach cancer and oesophageal cancer. The nucleic acids are useful for
XX detecting and monitoring cancer, particularly prostate cancer. They are
XX particularly useful in the treatment or prevention of cancer, in
XX producing DNA-based vaccines against prostate cancer or that promote
XX anti-tumor immune responses, and to raise antibodies. The expression of
XX genes and detection of their protein products and/or peptides may be
XX used to monitor disease progression during therapy, or as a prognostic
XX indicator of initial disease status of the patient.
XX
SQ Sequence 694 BP; 148 A; 153 C; 172 G; 203 T; 18 other;

Alignment Scores:
Pred. No.: 8.88e-06 Length: 694
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-000-039A-4 (1-16) x AAH42835 (1-694)

Oy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
Db 112 GAGGCTTTCCTAGGCTTTCTCTATGCGCTCCACGGACTCTTCTC 159

RESULT 2
AAF44639
ID AAF44639 standard; cDNA; 1296 BP.
XX
XX AAF44639;
XX
XX 27-MAR-2001 (first entry)
XX
XX Novel protein kinase cDNA, SEQ ID NO: 18.
XX
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX
XX Homo sapiens.
XX
XX WO200073469-A2.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 26-MAY-2000: 2000WO-US14842.
PF
XX

PR 28-MAY-1999: 99US-0136503.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
XX
XX WPI: 2001-032161/04.
DR P-PSDB; AAB65613.
XX
XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers
XX
XX Disclosure; Fig 2; 310pp; English.
XX
XX The present sequence encodes a novel protein kinase. The nucleic acids
XX and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
XX SQ Sequence 1296 BP; 345 A; 333 C; 293 G; 325 T; 0 other;

Alignment Scores:
Pred. No.: 1.87e-05 Length: 1296
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-10-000-039A-4 (1-16) x AAF44639 (1-1296)

Oy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
Db 1246 GAGGCTTTCCTAGGCTTTCTCTATGCGCTCCACGGACTCTTCTC 1293

RESULT 3
AAZ33647
ID AAZ33647 standard; cDNA; 2281 BP.
XX
XX AAZ33647;
XX
XX 08-DEC-1999 (first entry)
XX
XX Human breast tumour-associated EST 37.
XX
XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;
KW treatment; tumour; cytostatic; medicament; ss.
XX
XX Homo sapiens.
XX
XX DE19813839-A1.
PN
XX
XX 23-SEP-1999.
PD
XX
XX 20-MAR-1998; 98DE-1013839.
PF
XX
XX 20-MAR-1998; 98DE-1013839.
PR
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MHH.
PA

XX PI Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosentahl A;
 XX WPI: 1999-528981/45.
 DR P-PSDB; AAV48573.
 XX Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -
 XX Claim 1a; 116-117; 188pp; German.
 XX This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AAZ33611-248617 represents expressed
 CC sequence tags described in the method of the invention.
 XX SQ Sequence 2281 BP; 601 A; 498 C; 494 G; 688 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.66e-05 Length: 2281
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0
 US-10-000-039A-4 (1-16) x AAZ33647 (1-2281)
 QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 DB 1188 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGACTCTTCCTC 1235
 RESULT 4
 AAV23833
 ID AAV23833 standard; cDNA: 2311 BP.
 XX AC AAV23833;
 XX 31-JUL-1998 (first entry)
 XX Human protein kinase HPK-3 coding sequence.
 XX Protein kinase; human; HPK; signalling cascade; kinase expression;
 KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;
 KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;
 KW lymphoma; therapy; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 379..1080
 FT /*tag= a
 FT CDS
 XX WO9811234-A2.
 XX 19-MAR-1998.
 XX 10-SEP-1997; 97WO-US15923.
 XX 12-SEP-1996; 96US-0712709.
 XX (INCY-) INCYTE PHARM INC.
 XX Au-Young J, Guegler KJ, Hawkins PR;
 XX WPI: 1998-207394/18.
 DR P-PSDB; AAV54205.
 XX New isolated human protein kinase(s) - used to develop products for

PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
 PT asthma, multiple sclerosis or rheumatoid arthritis
 XX Claim 25; Fig 3; 75pp; English.
 XX This sequence encodes a human protein kinase (HPK) of the invention. The
 CC HPK protein can be used to develop products for studying signalling
 CC cascades in various cells and tissues, diagnosing disease and selecting
 CC inhibitors or drugs with the potential to intervene in various disorders
 CC or diseases in which altered kinase expression is implicated. The
 CC products can be used to e.g. reverse memory loss such as due to
 CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
 CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
 CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
 CC lymphocytic leukaemias or lymphomas.
 XX SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;
 Alignment Scores:
 Pred. No.: 3.71e-05 Length: 2311
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0
 US-10-000-039A-4 (1-16) x AAV23833 (1-2311)
 QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 DB 1255 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGACTCTTCCTC 1302
 RESULT 5
 ACA56174
 ID ACA56174 standard; cDNA: 2311 BP.
 XX AC ACA56174;
 XX 06-JUN-2003 (first entry)
 XX Human signalling pathway polynucleotide probe SEQ ID NO 772.
 XX Human; probe; ss; array element; Parkinson's disease;
 KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
 KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
 XX Homo sapiens.
 XX US6500938-B1.
 XX 31-DEC-2002.
 XX 30-JAN-1998; 98US-0016434.
 XX 30-JAN-1998; 98US-0016434.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Au-Young J, Seilhamer JJ;
 XX WPI: 2003-352189/33.
 XX Combination of polynucleotide probes, useful as array elements in a
 PT microarray for monitoring the expression of a number of target
 PT polynucleotides -
 XX Claim 1; SEQ ID NO 772; 65pp; English.
 XX The invention relates to a combination which, comprises a number of
 CC polynucleotide probes comprising a sequence selected from one of the 1490
 CC sequences mentioned in the specification. The combination is useful as an
 CC array element in a microarray for monitoring the expression of a number
 CC of target polynucleotides. The microarray is particularly useful in the

Tue Aug 12 10:56:07 2003

us-10-000-039a-4.rng

CC diagnosis and treatment of cancer and immunopathology and neuropathology.
 CC The microarray is useful in diagnostics and treatment regimens, drug
 CC discovery and development, toxicological and carcinogenicity studies,
 CC forensics and pharmacogenomics. The microarray is also useful for
 CC monitoring progression of diseases and for developing sophisticated
 CC profiles for the effects of currently available therapeutic drugs. The
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
 CC and genomic fragments and in research and diagnostic applications. The
 CC array can detect changes in expression in a large number of genes coding
 CC for different signaling pathway populations which can be used to diagnose
 CC various diseases including cancer e.g. adenocarcinoma and leukemia,
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a polynucleotide
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
 XX
 SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 592 T; 1 other;

Alignment Scores: 3.71e-05 Length: 2311
 Pred. No.: 85.00 Matches: 16
 Score: 100.00% Conservativity: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 25 Gaps: 0
 DB:

US-10-000-039A-4 (1-16) x ACA56174 (1-2311)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1255 GAGGCTTCTAGGCTTTCTATGCGCTCCACGAGCTCTTCTC 1302

RESULT 6
 ABX74395
 ID ABX74395 standard; cDNA; 2343 BP.
 XX
 AC ABX74395;
 XX
 DT 21-MAR-2003 (first entry)
 XX
 DE Human cDNA sequence #6 up-regulated in non-aggressive CC-RCC.
 XX
 KW Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200279411-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 29-MAR-2002; 2002WO-US09576.
 XX
 PR 29-MAR-2001; 2001US-279411P.
 XX
 PA (VAND-) VAN ANDEL INST.
 XX
 PI Haab B, Rhodes D, Teh BT, Takashi M;
 XX
 XX WPI: 2003-040679/03.

XX New microarray, comprising a matrix of cDNA probe from a set of probes
 XX immobilised to a solid surface in predetermined order, useful in the
 PT prognosis of patients with clear cell renal carcinoma
 PT
 XX Claim 1; Page 79-80; 17pp; English.
 PS
 XX The present invention relates to a microarray comprising a matrix of
 CC at least one cDNA probe from a set of probes immobilised to a solid
 CC

CC surface in a predetermined order, where a row of pixels corresponds
 CC to replicates of one distinct probe from the set. The probes are
 CC complementary to nucleic acid sequences that are expressed
 CC differentially in aggressive as compared to non-aggressive types of
 CC clear cell renal carcinoma (CC-RCC) and that hybridise to the probes
 CC under high stringency conditions. The microarray is useful for the
 CC prognosis of patients with CC-RCC, wherein aggressive and
 CC non-aggressive CC-RCC tumour types are characterised by differential
 CC expression profiles of genes that hybridise with one or more probes
 CC immobilised on the microarray. The arrays are useful for gene
 CC expression profiling of tumour and normal tissues. The present
 CC sequence represents a human cDNA sequence up-regulated in
 CC non-aggressive CC-RCC phenotypes.

SQ Sequence 2343 BP; 612 A; 517 C; 511 G; 703 T; 0 other;

Alignment Scores: 3.78e-05 Length: 2343
 Pred. No.: 85.00 Matches: 16
 Score: 100.00% Conservativity: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 25 Gaps: 0
 DB:

US-10-000-039A-4 (1-16) x ABX74395 (1-2343)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1281 GAGGCTTCTAGGCTTTCTATGCGCTCCACGAGCTCTTCTC 1328

RESULT 7
 AA74190
 ID AA74190 standard; DNA; 2346 BP.
 XX
 AC AA74190;
 XX
 DT 15-MAR-1999 (first entry)
 XX
 DE Human sgk DNA.
 XX
 KW Serum glucocorticoid regulated kinase; sgk; human; treatment; inhibitor;
 KW serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 KW chronic renal failure; inflammation; Alzheimers disease; wound; ss.
 XX
 OS Homo sapiens.
 XX
 FN Key Location/Qualifiers
 FT CDS 45..1340
 FT /*tag= "a"
 FT /product= "sgk"
 FT /transl_except= (pos:1185..1187, aa:Asp)
 XX
 EP887081-A2.
 XX
 PD 30-DEC-1998.
 XX
 PF 27-MAY-1998; 98EP-0304189.
 XX
 PR 27-JUN-1997; 97US-0051124.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Kumar JM;
 XX
 XX WPI: 1999-047627/05.
 DR P-PSDB; AAW90139.
 DR
 XX Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 PT disease - by administration of nucleic acids and antagonists which
 PT inhibit activity or expression of human serum glucocorticoid
 PT regulated kinase (sgk), a serine/threonine protein kinase
 XX
 PS Disclosure; Page 14-15; 17pp; English.

Tue Aug 12 10:56:07 2003

us-10-000-039a-4.rng

XX This sequence encodes a novel human serum glucocorticoid regulated kinase
 CC (sgk) protein which is a member of the serine/threonine protein kinase
 CC family. This protein is used for the treatment of a subject having need
 CC to inhibit/antagonise activity or expression of human sgk polypeptide
 CC e.g. for the treatment of chronic renal failure, diabetic nephropathy,
 CC inflammation, Alzheimers disease and wounds.
 XX Sequence 2346 BP: 608 A: 520 C: 518 G: 700 T: 0 other;
 Alignment Scores: Length: 2346
 Pred. No.: 3.78e-05 Matches: 16
 Score: 85.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 20 Gaps: 0
 DB: 1290 GAGGCTTCTAGGCTTTTCTATGCGCTCCACGACTCTTTCCTC 1337
 US-10-000-039A-4 (1-16) x AAV74190 (1-2346)
 QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 DB 1290 GAGGCTTCTAGGCTTTTCTATGCGCTCCACGACTCTTTCCTC 1337
 RESULT 8
 AAV48311
 ID AAV48311 standard; cDNA: 2370 BP.
 XX AC AAV48311;
 XX DT 16-NOV-1998 (first entry)
 XX DE Human cell-volume regulating kinase h-sgk.
 XX KW ss; human; cell-volume; kinase; h-sgk; diabetes mellitus;
 KW renal insufficiency; inflammation; Alzheimer's disease.
 XX OS Homo sapiens.
 XX FH Key
 FT CDS Location/Qualifiers
 FT CDS 43..1338
 FT /*tag= a
 FT /*product= "Kinase h-sgk"
 XX EP861896-A2.
 XX PD 02-SEP-1998.
 XX PF 27-JAN-1998; 98EP-0101338.
 XX PR 28-FEB-1997; 97DE-1008173.
 XX PA (DADE-) DADE BEHRING MARBURG GMBH.
 XX PI Lang F, Waldegger S;
 XX DR WPI: 1998-449109/39.
 XX DR P-PSDB; AAV77219.
 XX New nucleic acid encoding cell-volume regulating kinase h-sgk and
 PT related proteins - used for diagnosis and treatment of diseases
 PT involving changes in cell volume, e.g. renal insufficiency,
 PT inflammation, infections etc.
 XX Disclosure; Fig 1; 15pp; German.
 XX The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose

CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.
 XX Sequence 2370 BP: 636 A: 517 C: 513 G: 704 T: 0 other;
 Alignment Scores: Length: 2370
 Pred. No.: 3.83e-05 Matches: 16
 Score: 85.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 19 Gaps: 0
 DB: 1288 GAGGCTTCTAGGCTTTTCTATGCGCTCCACGACTCTTTCCTC 1335
 US-10-000-039A-4 (1-16) x AAV48311 (1-2370)
 QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
 DB 1288 GAGGCTTCTAGGCTTTTCTATGCGCTCCACGACTCTTTCCTC 1335
 RESULT 9
 ABK84749
 ID ABK84749 standard; cDNA: 2370 BP.
 XX AC ABK84749;
 XX DT 14-AUG-2002 (first entry)
 XX DE Human CDNA differentially expressed in granulocytic cells #1320.
 XX KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX OS Homo sapiens.
 XX PN WO200228999-A2.
 XX PD 11-APR-2002.
 XX PF 03-OCT-2001; 2001WO-US0821.
 XX PR 03-OCT-2000; 2000US-237189P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Beazer-Barclay Y, Weissman SM, Yamaya S, Vockley J;
 XX DR WPI: 2002-435328/46.
 XX Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity
 XX Claim 1; SEQ ID No 1320; 114pp; English.
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 CC (SCA), by detecting the level of expression of gene(s) (Gs) identified by
 CC DNA chip analysis as given in the specification, and comparing
 CC the expression level to an expression level in an unactivated
 CC GC, where differential expression of Gs is indicative of GCA.
 CC Also included are modulating (M2) GA by contacting GC with an agent
 CC that alters the expression of at least one gene in Gs; (2) screening (M3)
 CC for an agent capable of modulating GCA or an inflammation (especially
 CC chronic) in a tissue, an allergic response in a subject, exposure of a
 CC subject to a pathogen or sterile inflammatory disease using the
 CC gene expression profile; (3) detecting (M4) an inflammation (especially

chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, inflammatory disease (e.g. glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:
Pred. No.: 3.83e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-10-000-039a-4 (1-16) x ABK84749 (1-2370)

QY 1 GluAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 16

DB 1288 GAGGCTTTCTAGGCTTTTCCTATGCGCTCCCGAGACTCTTTCCTC 1335

RESULT 10

ABN97360

ID ABN97360 standard; DNA: 2370 BP.

XX AC ABN97360;

XX 13-AUG-2002 (first entry)

XX Gene #3858 used to diagnose liver cancer.

XX Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley Jg;

XX WPI; 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient, in a
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample

PS Claim 1; SEQ ID NO 3858; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumour in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:

Pred. No.: 3.83e-05 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-10-000-039a-4 (1-16) x ABN97360 (1-2370)

QY 1 GluAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 16

DB 1288 GAGGCTTTCTAGGCTTTTCCTATGCGCTCCCGAGACTCTTTCCTC 1335

RESULT 11

ABL70006

ID ABL70006 standard; DNA: 2370 BP.

XX AC ABL70006;

XX 15-MAY-2002 (first entry)

XX Pancreas cancer related gene sequence SEQ ID NO:8343.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233133P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 20-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 22-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 29-SEP-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PA (AVAL-) AVALON PHARM.
XX
XX
FI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1: SEQ ID 8343; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in AB161664
XX to AB170110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 3.83e-05 Length: 2370
XX Score: 85.00 Matches: 16
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-10-000-039A-4 (1-16) x AB170006 (1-2370)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
DB 1288 GAGGCTTCTAGGCTTTCTCTATCGCGCTCCACGAGACTCTTCTC 1335
RESULT 12
ABI99409
ID ABI99409 standard; cDNA: 431 BP.
XX
AC ABI99409;
XX
DT 07-MAR-2002 (first entry)
XX
XX Mouse ischaemic condition related cDNA sequence SEQ ID NO:325.
DE
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
XX Mus musculus.
OS
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-JP04192.
PF
XX 18-MAY-2000; 2000JP-0145977.
PR
XX (UYNI-) UNIV NITHON SCHOOL JURIDICAL PERSON.
PA
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
XX Claim 2: Page 893; 2690pp; English.
XX
XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (ABI99202 to AB199912, encoding
XX the protein sequences in AB857020 to AB857374) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischaemic condition-improving
XX drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
XX represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
XX Sequence 431 BP; 75 A; 128 C; 115 G; 111 T; 2 other;
XX
XX Alignment Scores:
XX Pred. No.: 3.9e-05 Length: 431
XX Score: 80.00 Matches: 15
XX Percent Similarity: 93.75% Conservative: 0
XX Best Local Similarity: 93.75% Mismatches: 1
XX Query Match: 94.12% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-10-000-039A-4 (1-16) x ABI99409 (1-431)
XX
QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
DB 105 GAAGCCTTCTAGGCTTTCTCTATGACCTCTCTGTGGATTCTTCTC 152
RESULT 13

Tue Aug 12 10:56:07 2003

us-10-000-039a-4.rng

AA44640
ID AAF44640 standard; cDNA: 2432 BP.
XX
AC AAF44640;
XX
DT 27-MAR-2001 (first entry)
XX
DE Novel protein kinase cDNA, SEQ ID NO: 19.
XX
KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX
OS Mus musculus.
XX
PN WO200073469-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000: 2000WO-US14842.
XX
PR 28-MAY-1999: 99US-0136503.
XX
PA (SUGEN) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
DR WPI: 2001-032161/04.
DR P-PSDB; AAB65614.
XX
PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -
XX
PS Disclosure; Fig 2; 310pp; English.
XX
CC The present sequence encodes a novel protein kinase. The nucleic acids
CC and the protein kinases they encode may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
CC reproductive disorders.
XX
SQ Sequence 2432 BP; 586 A; 586 C; 560 G; 699 T; 1 other;

Alignment Scores:
Pred. No.: 0.000305 Length: 2432
Score: 80.00 Matches: 15
Percent Similarity: 93.75% Conservative: 0
Best Local Similarity: 93.75% Mismatches: 1
Query Match: 94.12% Indels: 0
DB: 22 Gaps: 0
US-10-000-039A-4 (1-16) x AAF44640 (1-2432)
Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspPheLeu 16
Db 1317 GAAGCCTTCCTGGGCTTCCTGATGACCTCTCTGGATTCCTCTC 1364

RESULT 14
ABX42261
ID ABX42261 standard; cDNA: 404 BP.
XX
AC ABX42261;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #7426.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-0960352.
XX
PR 12-JAN-1999: 99US-115707P.
PR 11-JAN-2000; 2000US-0480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
DR WPI: 2003-110599/10.
XX
CC New nucleic acid associated with lactation, and muscle and fat
CC deposition, useful for genome mapping, gene identification and
CC analysis, cattle breeding, or for genetically improving cattle
XX
PS Claim 2; SEQ ID No 7426; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived
CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
CC second nucleic acid molecule comprising any of 15112 nucleotide
CC sequences, appearing as ABX34836-ABX49847, or complements of them.
CC Also included are: (1) a transformed cell having a nucleic acid
CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
CC translated sequence that functions in the cell to cause termination of
CC transcription and addition of polyadenylated ribonucleotides to a 3'
CC of the mRNA molecule; and (2) determining a level or pattern of a
CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
CC complement or fragment with a complementary nucleic acid sequence or its
CC complement or fragment with a complementary nucleic acid molecule
CC obtained from the bovine cell or tissue, where hybridisation between the
CC marker nucleic acid and the complementary nucleic acid permits the
CC detection of the molecule; and (b) detecting the level or pattern of the
CC complementary nucleic acid, where the detection of the complementary
CC nucleic acid is predictive of the level or pattern of the molecule.
CC The LMFD nucleic acid is used for determining a level or pattern
CC of a molecule in a bovine cell or tissue. It is useful for genome
CC mapping, gene identification and analysis, cattle breeding, preparation
CC of constructs for use in cattle gene expression, or for genetically
CC improving cattle. The present sequence is one of the 15112 bovine
CC LMFD EST (expressed sequence tag) nucleic acids.
CC Note: The present sequence was not shown in the specification but
CC was obtained in electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocId=20020137139.
XX
SQ Sequence 404 BP; 97 A; 108 C; 101 G; 98 T; 0 other;

Alignment Scores:
Pred. No.: 5.43e-05 Length: 404
Score: 79.00 Matches: 15

The present sequence encodes a rat serum and glucocorticoid induced protein kinase (SGK). The specification describes a screening method to identify a compound which modulates the activity of protein kinases from different sources, using host yeast cells. The method is used to identify a compound which modulates (inhibits) the activity of a protein kinase. Pkh1 or Pkh2 phosphorylate and activate Ypk1, Yrk2, SGK or protein kinase B-alpha (PKBa/pha). 3-phosphoinositide-dependent protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and Yrk2 or SGK but not PKBa/pha or p70S6 kinase. Compounds identified by the methods are used to treat fungal infections e.g. thrush, and to treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be used in the treatment of diabetes or obesity, and compounds which inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK (Ypk1 or Yrk2) can be used as antifungal agent to treat Candida

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:36:34 ; Search time 15.2727 Seconds
(without alignments)
462.401 Million cell updates/sec

Title: US-10-000-039a-4

Perfect score: 85

Sequence: 1 EAFGLSYAPPTDSFL 16

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlp
-Q=/cgn2_1/USPRO-spool/US10000039/runat_01082003.085419.27535/app_query.fasta.1.661
-DB-Issued Patents_NA -QFWT-fastcap -SUFFIX=rni -MINWATCH=0.1 -LOOPFCI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039 -CGCN_1_1_145_runat_01082003.085419.27535 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTOUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	2311	2	US-08-712-709-6
2	85	100.0	2311	3	US-09-111-444-6
3	85	100.0	2311	3	US-09-541-228-6
4	85	100.0	2311	4	US-09-016-434-772
5	85	100.0	2370	4	US-09-031-295-1
6	46	54.1	1440	4	US-09-620-312D-614
7	45	52.9	993	4	US-09-252-991A-7223
8	45	52.9	38584	4	US-09-453-702B-50
9	45	52.9	1230025	4	US-09-198-452A-1
10	43	50.6	1607	2	US-08-749-902-4
11	43	50.6	1607	4	US-09-016-434-126
12	43	50.6	1732	4	US-09-430-564-1

c	13	42	49.4	2281	4	US-09-392-772-9	Sequence 9, Appl 1
c	14	42	49.4	5197	1	US-08-131-365B-53	Sequence 5, Appl 1
c	15	42	49.4	5197	2	US-08-668-123-53	Sequence 5, Appl 1
c	16	42	49.4	5332	4	US-09-801-861-3	Sequence 3, Appl 1
c	17	41	48.2	489	1	US-08-663-310-3	Sequence 3, Appl 1
c	18	41	48.2	489	1	US-08-663-310-10	Sequence 10, Appl 1
c	19	41	48.2	489	2	US-09-006-491-3	Sequence 3, Appl 1
c	20	41	48.2	489	2	US-09-006-491-10	Sequence 10, Appl 1
c	21	41	48.2	489	3	US-09-335-919-3	Sequence 3, Appl 1
c	22	41	48.2	489	3	US-09-335-919-10	Sequence 10, Appl 1
c	23	41	48.2	1261	3	US-08-980-832-30	Sequence 30, Appl 1
c	24	41	48.2	1631	1	US-08-663-310-12	Sequence 12, Appl 1
c	25	41	48.2	1631	2	US-09-006-491-12	Sequence 12, Appl 1
c	26	41	48.2	1631	3	US-09-335-919-12	Sequence 12, Appl 1
c	27	41	48.2	1741	3	US-08-795-430-12	Sequence 12, Appl 1
c	28	41	48.2	1741	4	US-09-355-700-12	Sequence 12, Appl 1
c	29	41	48.2	1744	1	US-08-424-406-7	Sequence 7, Appl 1
c	30	41	48.2	1744	1	US-08-464-233B-7	Sequence 7, Appl 1
c	31	41	48.2	2886	1	US-08-663-310-7	Sequence 7, Appl 1
c	32	41	48.2	2886	2	US-09-006-491-7	Sequence 7, Appl 1
c	33	41	48.2	2886	3	US-09-335-919-7	Sequence 7, Appl 1
c	34	41	48.2	14255	1	US-08-320-559-1	Sequence 1, Appl 1
c	35	41	48.2	14255	1	US-08-327-392-1	Sequence 1, Appl 1
c	36	41	48.2	14255	1	US-08-306-691B-55	Sequence 55, Appl 1
c	37	41	48.2	14255	3	US-08-545-860D-1	Sequence 1, Appl 1
c	38	41	48.2	14255	5	PCT-US94-04496-1	Sequence 1, Appl 1
c	39	41	48.2	62804	4	US-09-800-960-3	Sequence 3, Appl 1
c	40	40.5	47.6	1176	4	US-09-009-913-1	Sequence 1, Appl 1
c	41	40.5	47.6	1176	4	US-09-134-001C-853	Sequence 853, App
c	42	40	47.1	204	4	US-09-252-991A-13094	Sequence 13094, A
c	43	40	47.1	278	3	US-09-328-111-390	Sequence 390, App
c	44	40	47.1	369	4	US-09-252-991A-12545	Sequence 12545, A
c	45	40	47.1	606	3	US-09-385-982-295	Sequence 295, App

ALIGNMENTS

RESULT 1
US-08-712-709-6
; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PR-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6

Alignment Scores:
Pred. No.: 4.12e-06 Length: 2311
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-000-039A-4 (1-16) x US-08-712-709-6 (1-2311)
Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
Db 1255 GAGGCTTCCTAGGCTTTTCCTATGCGCTCCCGGACTCTTCTC 1302

RESULT 2
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-09-111-444-6

Alignment Scores:
Pred. No.: 4.12e-06 Length: 2311
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-000-039A-4 (1-16) x US-09-111-444-6 (1-2311)
Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
Db 1255 GAGGCTTCCTAGGCTTTTCCTATGCGCTCCCGGACTCTTCTC 1302

RESULT 3
; Sequence 6, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip K.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/541,228
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-09-541-228-6

Alignment Scores:
Pred. No.: 4.12e-06 Length: 2311
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-000-039A-4 (1-16) x US-09-541-228-6 (1-2311)
Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16
Db 1255 GAGGCTTCCTAGGCTTTTCCTATGCGCTCCCGGACTCTTCTC 1302

RESULT 4
US-09-016-434-772
```

; Sequence 772, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; CLASSIFICATION:
; NAME: Zeiler, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
; US-09-016-434-772

Alignment Scores:
Pred. No.: 4,12e-06 Length: 2311
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-000-039a-4 (1-16) x US-09-016-434-772 (1-2311)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
|||||
Db 1255 GAGGCTTTCCTAGGCTTTTCCTATGGCGCTCCACGGACTCTTCTC 1302

RESULT 5
US-09-031-295-1
; Sequence 1, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGER, Tubingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington

; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..1335
; US-09-031-295-1

Alignment Scores:
Pred. No.: 4,26e-06 Length: 2370
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-000-039a-4 (1-16) x US-09-031-295-1 (1-2370)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 16
|||||
Db 1288 GAGGCTTTCCTAGGCTTTTCCTATGGCGCTCCACGGACTCTTCTC 1445

RESULT 6
US-09-620-312D-614
; Sequence 614, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B

REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 126:
SEQUENCE CHARACTERISTICS:
LENGTH: 1607 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLAFET02
CLONE: 1309709
US-09-016-434-126

Alignment Scores:
Pred. No.: 131 Length: 1607
Score: 43.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 50.59% Indels: 0
DB: 4 Gaps: 0

US-10-000-039A-4 (1-16) x US-09-016-434-126 (1-1607)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThr 12
Db 1253 CAGGCCCTCTGGGCTTCACATACGTCGCGCGCT 1288

RESULT 12

US-09-430-564-1
Sequence 1, Application US/09430564
Patent No. 6372467
GENERAL INFORMATION:
APPLICANT: John Blenis
APPLICANT: Kay K. Lee-Fruman
APPLICANT: Calvin J. Kuo
TITLE OF INVENTION: P5456K AND P8556K GENES, PROTEINS,
FILE REFERENCE: 00246/506002
CURRENT APPLICATION NUMBER: US/09/430.564
CURRENT FILING DATE: 1999-10-29
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1732
TYPE: DNA
ORGANISM: Homo sapiens
US-09-430-564-1

Alignment Scores:
Pred. No.: 143 Length: 1732
Score: 43.00 Matches: 7
Percent Similarity: 83.33% Conservative: 3
Best Local Similarity: 58.33% Mismatches: 2
Query Match: 50.59% Indels: 0
DB: 4 Gaps: 0

US-10-000-039A-4 (1-16) x US-09-430-564-1 (1-1732)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThr 12
Db 1179 CAGGCCCTCTGGGCTTCACATACGTCGCGCGCT 1214

RESULT 13

US-09-392-772-9/c
Sequence 9, Application US/09392772
Patent No. 6346403
GENERAL INFORMATION:
APPLICANT: Famodu, Layo O.

APPLICANT: Antoni Rafalski
TITLE OF INVENTION: Methionine Metabolic Enzymes
REFERENCE: BB-1241
CURRENT APPLICATION NUMBER: US/09/392.772
CURRENT FILING DATE: 1999-09-07
EARLIER APPLICATION NUMBER: 60/099,519
EARLIER FILING DATE: 1998-09-08
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Microsoft Office 97
SEQ ID NO 9
LENGTH: 2281
TYPE: DNA
ORGANISM: zea mays
US-09-392-772-9

Alignment Scores:
Pred. No.: 307 Length: 2281
Score: 42.00 Matches: 6
Percent Similarity: 68.75% Conservative: 5
Best Local Similarity: 37.50% Mismatches: 5
Query Match: 49.41% Indels: 0
DB: 4 Gaps: 0

US-10-000-039A-4 (1-16) x US-09-392-772-9 (1-2281)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheIeu 16
Db 1452 GATGTTCTTCGCTCTTTTGTGTCCTCCATCAGATATTTTGTG 1405

RESULT 14

US-08-131-365B-53
Sequence 53, Application US/08131365B
Patent No. 5527690
GENERAL INFORMATION:
APPLICANT: Brown, Michael S.
APPLICANT: Briggs, Michael R.
APPLICANT: Wang, Xiaodong
APPLICANT: Goldstein, Joseph L.
TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING
TITLE OF INVENTION: TO STEROL REGULATORY ELEMENT BINDING
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Burke
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/131,365B
FILING DATE: 01-OCT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: OTSD:372/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 5197 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"

US-08-131-365B-53

Alignment Scores: 852 Length: 5197
Pred. No.: 42.00 Matches: 10
Score: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 5
Query Match: 49.41% Indels: 0
DB: 1 Gaps: 0

US-10-000-039A-4 (1-16) x US-08-131-365B-53 (1-5197)

Qy 2 AlaPheLeuGlyPheSerTyAlaProThrAspSerPheLeu 16
Db 3838 GCCTTCTGAGTTTCTCTCTGACCCCTACTCTCTCTTTTG 3882

RESULT 15

US-08-668-123-53
; Sequence 53, Application US/08668123
; Patent NO. 5891631
; GENERAL INFORMATION:
; APPLICANT: Brown, Michael S.
; APPLICANT: Briggs, Michael R.
; APPLICANT: Wang, Xiaodong
; APPLICANT: Goldstein, Joseph L.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING
; TO STEROL REGULATORY ELEMENT BINDING
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.A.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,123
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/131,365
; FILING DATE: 01-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:372/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5197 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"

US-08-668-123-53
Alignment Scores: 852 Length: 5197
Pred. No.: 42.00 Matches: 10
Score: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 5
Query Match: 49.41% Indels: 0
DB: 2 Gaps: 0

US-10-000-039A-4 (1-16) x US-08-668-123-53 (1-5197)

Qy 2 AlaPheLeuGlyPheSerTyAlaProThrAspSerPheLeu 16
Db 3838 GCCTTCTGAGTTTCTCTCTGACCCCTACTCTCTCTTTTG 3882

Search completed: August 11, 2003, 12:08:51
Job time : 116.273 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 07:58:29 ; Search time 115.612 Seconds
(without alignments)
962.020 Million cell updates/sec

Title: US-10-000-039A-2
Perfect score: 2270
Sequence: 1 MTWKTEAAKGLTYSRMGM.....KEAAEAFGLFSYAPPTDSFL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp_unclassified.*
- 15: sp-rvivirus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2071.5	91.3	434	13	093524 xenopus lae
2	1885	83.0	433	13	073926 squalus aca
3	1818	80.1	594	13	073927 squalus aca
4	1480	65.2	429	11	08VEK1
5	1072	47.2	422	5	Q94365 caenorhabdi
6	947	41.7	185	11	Q99LU4
7	905.5	39.9	486	5	Q95XJ0
8	890	39.2	530	5	Q24293
9	890	39.2	530	5	Q879A5
10	890	39.2	611	5	Q24469
11	890	39.2	611	5	Q81NB9
12	872	38.4	479	13	Q8UUX0
13	865	38.1	480	13	Q57513
14	863.5	38.0	481	13	Q98TY9
15	851	37.5	452	11	Q8BS26
16	850	37.4	480	4	Q95WB6

Q96vr1 cryphonectr
Q8ce74 mus musculus
Q99012 trichoderma
Q9pfj8 schizosacch
Q17942 caenorhabdi
Q9brs0 homo sapien
Q9w6y9 xenopus lae
Q8chx0 mus musculus
Q8bsk8 mus musculus
Q8jtz9 lugu rubrip
Q17941 caenorhabdi
Q9mgs0 botrytis ci
Q9mxb7 aplysia cal
Q9gqb3 artemia sal
Q95t78 drosophila
Q95t78 drosophila
Q94533 drosophila
P91656 drosophila
Q20953 caenorhabdi
Q8nik8 emericeella
Q814w3 plasmodium
P90980 caenorhabdi
Q8mg88 caenorhabdi
Q62567 suberites d
O01715 hydra atten
O01716 hydra atten
Q9xtg7 caenorhabdi
Q8mb66 limulus pol
Q8sqh7 entamoeba h
Q90xf2 brachydanio

ALIGNMENTS

RESULT 1

093524 ID O93524 PRELIMINARY; PRT: 434 AA.
AC O93524;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Hypothetical 49.1 kDa protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=renal tubule;
RX MEDLINE=99162637; PubMed=10051674;
RA Chen S.-Y., Bhargava A., Mastrobardino L., Meijer O.C., Wang J.,
Buse P., Firestone G.L., Verrey F., Pearce D.,
RT "Epithelial sodium channel regulated by aldosterone-induced protein
sgk.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2514-2519(1999).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF057138; AAC62398.1; -
DR HSSP; P00517; 1YDR.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR SMART; SM000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferrase.
SQ SEQUENCE 434 AA; 49130 MW; 4A061E38B6AA6F61 CRC64;

```

Query Match      91.3%; Score 2071.5; DB 13; Length 434;
Best Local Similarity 89.9%; Pred. No. 1.7e-164;
Matches 390; Conservative 25; Mismatches 16; Indels 3; Gaps 2;

QY 1 MTVKTEAAGK--TLTYSRMRGMVAIIIAFMKQRRMGLNDFIQKIA-NNSYACKHPSEVOSI 57
DB 1 MTVKTEAAGK--TLTYSRMRGMVAIIIAFMKQRRMGLNDFIQKIA-NNSYACKHPSEVOSI 60
QY 58 LKISQPEPELMNANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARH 117
DB 61 LKISQPEPELMNANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARH 120
QY 118 KAEVEFYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFVLGLHFSFQTAADKLYFVLD 177
DB 121 QSEKEFYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFVLGLHFSFQTAADKLYFVLD 180
QY 178 YINGGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIV 237
DB 181 YINGGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIV 240
QY 238 LTFDGLCKENIEHNSTTFCGTPPEYLAPEVLHKQPYDRTVDMWCLGAVLYEMLYGLPPF 297
DB 241 LTFDGLCKENIEHNSTTFCGTPPEYLAPEVLHKQPYDRTVDMWCLGAVLYEMLYGLPPF 300
QY 298 YSRNTAEYDNIILNKLPLQKPNITNSARHLLGLLQKDRKRLGAKDDFMEIKSHVFFSL 357
DB 301 YSRNTAEYDNIILNKLPLQKPNITNSARHLLGLLQKDRKRLGAKDDFMEIKSHVFFSL 360
QY 358 INWDLLINKKITPPENPNVSGPNELRHFDPEETPEEPVNSIGKSPDSVLVATASVKEAEEA 417
DB 361 INWDLLINKKITPPENPNVSGPNELRHFDPEETPEEPVNSIGKSPDSVLVATASVKEAEEA 420
QY 418 FLGFSYAPPTDSFL 431
DB 421 FMGFSYAPPMESYL 434

RESULT 2
O73926 PRELIMINARY; PRT; 433 AA.
AC O73926;
DT 01-AUG-1998 (TREMREL. 07, Created)
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE S-99ki.
OS Squalus acanthias (spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RA Waldegger S., Barth P., Forrest J., Greger R.F., Lang F.;
RT "Cloning of sgk Serine-Threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ223715; CAAL1527.1;
DR HSSP; P05132; LFMO.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TKC_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 433 AA; 48369 MW; 982BD01A59AA3A78 CRC64;

Query Match      83.0%; Score 1885; DB 13; Length 433;
Best Local Similarity 83.2%; Pred. No. 6.4e-149;
Matches 361; Conservative 33; Mismatches 36; Indels 4; Gaps 4;

QY 1 MTVKTEAAGKTLTYSRMRGMVAIIIAFMKQRRMGLNDFIQKI-ANNSYACKH-PEVQSIL 58
DB 1 MTIKSSSSGSAITYSKMRGVVAIIIAFMKQRRMGLNDFIQKLAASQSYACKHSSEVPAMI 60
QY 59 KISQPEPELMNANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARH 117
DB 61 HIS-PAETENGTGASPPSPPTTAQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARH 119
QY 118 KAEVEFYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFVLGLHFSFQTAADKLYFVLD 177
DB 120 KADQDFYAVKVLQKAILKKKEKHIMSERNVLLKNVHPFVLGLHFSFQTAADKLYFVLD 179
QY 178 YINGGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIV 237
DB 180 YINGGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLPENILLDSQGHIV 239
QY 238 LTFDGLCKENIEHNSTTFCGTPPEYLAPEVLHKQPYDRTVDMWCLGAVLYEMLYGLPPF 297
DB 240 LTFDGLCKENIEHNSTTFCGTPPEYLAPEVLHKQPYDRTVDMWCLGAVLYEMLYGLPPF 299
QY 298 YSRNTAEYDNIILNKLPLQKPNITNSARHLLGLLQKDRKRLGAKDDFMEIKSHVFFSL 357
DB 300 YSRNTAEYDNIILNKLPLQKPNITNSARHLLGLLQKDRKRLGAKDDFMEIKSHVFFSL 359
QY 358 INWDLLINKKITPPENPNVSGPNELRHFDPEETPEEPVNSIGKSPDSVLVATASVKEAEEA 417
DB 360 INWDLLINKKITPPENPNVSGPNELRHFDPEETPEEPVNSIGKSPDSVLVATASVKEAEEA 419
QY 418 FLGFSYAPPTDSFL 431
DB 420 FLGFSYAPPMESYL 433

RESULT 3
O73927 PRELIMINARY; PRT; 594 AA.
AC O73927;
DT 01-AUG-1998 (TREMREL. 07, Created)
DT 01-AUG-1998 (TREMREL. 07, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE S-99ki.
OS Squalus acanthias (spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RA Waldegger S., Barth P., Forrest J., Greger R.F., Lang F.;
RT "Cloning of sgk Serine-Threonine Protein Kinase from Shark Rectal
RT Gland - a Gene Induced by Hypertonicity and Secretagogues."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ223716; CAAL1528.1;
DR HSSP; P00517; 1YDR.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TKC_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 594 AA; 66857 MW; 40153EC18057B767 CRC64;

Query Match      80.1%; Score 1818; DB 13; Length 594;

```

Best Local Similarity 85.1%; Pred. No. 3.8e-143;

Matches 348; Conservative 26; Mismatches 31; Indels 4; Gaps 4;

```

QY 26 AFMKRRMGLNDFIOKI-ANNSYACKH-PEVOSILKISQOPEPELMNANSPSPSP-SCQ 82
Db 187 AFMKRRMGLNDFIOKLAASQSTACKSHSEVAMLIHS-PTETENGTGASPPSPPTAQ 245
QY 83 INLGSSNPHAKPSDFHFLKVIKGSFGKVLARHAKAEVEFYAVKVLQKKAIIKKKEKH 142
Db 246 INLGSSNPHAKPSDFHFLKVIKGSFGKVLARHAKADDQFYAVKVLQKKAIIKKKEKH 305
QY 143 INSERNVLLKNVKKHFLVGLHFSFQADKLYFVLDYINGGELFYHLQRCFCLEPRARY 202
Db 306 INSERNVLLKNVKKHFLVGLHFSFQADKLYFVLDYINGGELFYHLQRCFCLEPRARY 365
QY 203 AAEIASALGYLHSLNIVYRDLPENILLDSQGHVLTDFGLCKENIEHNSTTFTCGTPE 262
Db 366 AAEIASALGYLHSLNIVYRDLPENILLDRQGHVLTDFGLCKENIEHNSTTFTCGTPE 425
QY 263 YLAPEVLHKQPYDRTVDWMCGLGAVLYEMLYGLPPFYSRNTAEYDNLNKLQKPNITN 322
Db 426 YLAPEVLHKQPYDRTVDWMCGLGAVLYEMLYGLPPFYSRNTAEYDNLNKLQKPNISN 485
QY 323 SARHLEGLLOKDRKRLGAKDDFMEIKSHVFFSLINWDDLINKKITPPFNPNVSGPNEL 382
Db 486 SARHLEGLLOKDRKRLGAKDDFMEIKSHVFFSLINWDDLINKKITPPFNPNVSGPNEL 545
QY 383 RHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGLFSYAPPTDSFL 431
Db 546 QHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGLFSYAPPTDSFL 594

```

RESULT 4

```

Q8VEK1 ID Q8VEK1 PRELIMINARY; PRT; 429 AA.
AC Q8VEK1
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Similar to serum/glucocorticoid regulated kinase-like.
GN SGK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; BC018363; AAH18363.1; -.
DR MGD; MGI:2182368; Sgk3.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001683; PX.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00059; pkinase.1.
DR Pfam; PF00433; pkinase_C.1.
DR Pfam; PF00787; PX; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00195; PX; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 429 AA; 49077 MW; A9C6B7A0C34031F3 CRC64;

```

Query Match

Best Local Similarity 65.2%; Score 1480; DB 11; Length 429;

Matches 280; Conservative 54; Mismatches 70; Indels 8; Gaps 3;

```

QY 27 FMKORRMGLNDFIOKIANNSSYACKHPEVOSILKI-----SQOPEPELMNANSPSPSPS 80
Db 19 FIKORRAGLNEIQNLVRYPELYNHPDVNRAFLQMSDPRHQSDPSEDE-DEKSTSKPHSTYS 77
QY 81 QOINLGSSNPHAKPSDFHFLKVIKGSFGKVLARHAKAEVEFYAVKVLQKKAIIKKKEE 140
Db 78 RNINLGPTGNPHAKPTDFDLKVIKGSFGKVLAKRKLQKGFYAVKVLQKKAIIKKKEEQ 137
QY 141 KHMSERNVLLKNVKKHFLVGLHFSFQADKLYFVLDYINGGELFYHLQRCFCLEPRAR 200
Db 138 KHMAERNVLLKNVKKHFLVGLHFSFQADKLYFVLDYINGGELFYHLQRCFCLEPRAR 197
QY 201 FYAAEIASALGYLHSLNIVYRDLPENILLDSQGHVLTDFGLCKENIEHNSTTFTCGT 260
Db 198 FYAAEIASALGYLHSLNIVYRDLPENILLDSQGHVLTDFGLCKENIEHNSTTFTCGT 257
QY 261 PEYLAPEVLHKQPYDRTVDWMCGLGAVLYEMLYGLPPFYSRNTAEYDNLNKLQKPNITN 320
Db 258 PEYLAPEVLHKQPYDRTVDWMCGLGAVLYEMLYGLPPFYSRNTAEYDNLNKLQKPNITN 317
QY 321 TNSARHLEGLLOKDRKRLGAKDDFMEIKSHVFFSLINWDDLINKKITPPFNPNVSGPN 380
Db 318 SUTANSIIEELLEKKRQRLGAKDDFLEIONHPFFESLSWTDLVOKKIPPPNPNVAGPD 377
QY 381 ELRHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGLFSYAPPTDSFL 431
Db 378 DIRNFDVFTETVPVNSVCSVDYISVNASVLEADAFVGSFYAPPSDFL 429

```

RESULT 5

```

Q94365 ID Q94365 PRELIMINARY; PRT; 422 AA.
AC Q94365;
DT 01-FEB-1997 (TremBLrel. 02, Created)
DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
DE 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE W1066.2 protein.
GN W1066.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
OC Rhabditidae; Felodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT *Genome sequence of the nematode C.elegans: A platform for investigating biology.
RL Science 282:2012-2018(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z81140; CAB03485.1; -.
DR HSSP; P05132; IATP.
DR WormPep; W1066.2; CEI4798.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00059; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 422 AA; 49672 MW; 1FA75FF677B88C00 CRC64;

```

Query Match

Best Local Similarity 47.2%; Score 1072; DB 5; Length 422;

Matches 280; Conservative 54; Mismatches 70; Indels 8; Gaps 3;

Matches	201	Conservative	70	Mismatches	94	Indels	18	Gaps	4
QY	27	FMKORRMGLNDFQKIANNSYACKHPEVQSIKISQPOPELMNANPSPSPSOQINLG	86						
DB	38	FYEKRRWLVISQHLVDNN--LRSEDVRRFFHLESPODDE-----NNVDLG	82						
QY	87	PSNPHAKPSDFHFLKVIKQSGPKVLLARHAEVYAVKVLQKALKKKKEKHIMSE	146						
DB	83	PSERKTATANDFDYLTIGKSGFRVYQVHRKTKIYAKIILSKHEHIRKKNVEHVMAE	142						
QY	147	RNVLLKNVHPFLVGLRHFQFADKLYFVLVDVINGGELFYHLQRECFLEPRAREYAAEI	206						
DB	143	RNVLLKNVHPFLVGLRHFQFADKLYFVLVDVINGGELFYHLQRECFLEPRAREYAAEI	202						
QY	207	ASALGYLHNSIVYRDLKPNILDSQGHIVLTDGLCKENIEHNSSTFGCTGPEYLA	266						
DB	203	ACALGYLHNSIVYRDLKPNILDSQGHIVLTDGLCKENIEHNSSTFGCTGPEYLA	262						
QY	267	EVLHKKOPDRTVDMWCLGAVLYEMLYGLPPFSRNTAEYDNLNKLKPNITNSARH	326						
DB	263	EVLHKKOPDRTVDMWCLGAVLYEMLYGLPPFSRNTAEYDNLNKLKPNITNSARH	322						
QY	327	LLEGLQKTKRLGAKDPMFKSHVFFSLINWDDLKKTTPPPNPNVSGPNELRHPD	386						
DB	323	LITGLQKTKRLGAKDPMFKSHVFFSLINWDDLKKTTPPPNPNVSGPNELRHPD	382						
QY	387	PEFTEPEV-PNSIGKSPDSVLVT	408						
DB	383	KEFVEIQIDPSSL--NQQLAVT	403						
RESULT 6									
ID	Q99LU4	PRELIMINARY;	PRT;	185	AA.				
AC	Q99LU4;								
DT	01-JUN-2001 (TrEMBLrel. 17, Created)								
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)								
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)								
DE	Similar to serum/glucocorticoid regulated kinase (Fragment).								
GN	SGK.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID-10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Strausberg R.;								
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; BC002222; AA002222.1;								
DR	MGD; MGI:1340062; Sgk.								
DR	InterPro; IPR000961; Pkinase.C.								
DR	InterPro; IPR000719; Prot_kinase.								
DR	Pfam; PF00069; Pkinase.1.								
DR	Pfam; PF00433; Pkinase.C.1.								
DR	ProDom; PD000001; Prot_kinase.1.								
DR	PROSITE; PS00133; S_TK_X.1.								
DR	PROSITE; PS00101; PROTEIN_KINASE_DOM.1.								
KW	ATP-binding; Kinase; Transferase.								
FT	NON_TER								
SQ	SEQUENCE 185 AA; 2021 MW; EDAEA44ABF083945 CRC64;								
Query Match 41.7%; Score 947; DB 11; Length 185;									
Best Local Similarity 95.6%; Pred. No. 3.9e-71;									
Matches 173; Conservative 6; Mismatches 2; Indels 0; Gaps 0;									
QY	251	NSTSTFGCTGPEYLAPEVHLKQPDYDRTVDMWCLGAVLYEMLYGLPPFSRNTAEYDNL	310						
DB	5	NGTSTFGCTGPEYLAPEVHLKQPDYDRTVDMWCLGAVLYEMLYGLPPFSRNTAEYDNL	64						
QY	311	NKPLQKPNITNSARHLEGLQKDRTKRLGAKDPMFKSHVFFSLINWDDLKKTTP	370						
DB	65	NKPLQKPNITNSARHLEGLQKDRTKRLGAKDPMFKSHVFFSLINWDDLKKTTP	124						
Query Match 39.9%; Score 905.5; DB 5; Length 486;									
Best Local Similarity 50.0%; Pred. No. 4.1e-67;									
Matches 186; Conservative 49; Mismatches 114; Indels 23; Gaps 6;									
QY	59	KISQPOPELMNANPSPSPSQINLCPSPNPAKP--SDPFHLKVLKVGKSGFKVILARH	117						
DB	125	RVSCTNEEKTAQONARITDSKM-----PKRHTKTLQDEEFKMLGKGTFCCKVLLVRE	177						
QY	118	RAEYFYAVKVLQKALKKKKEKHIMSEVNLNKLKPNITNSARHLEGLQKDRTKRLGAKDPMFKSHVFFSLINWDDLKKTTP	177						
DB	178	KTGGELYATKILKRAVIVAKDEVARTLTESHVLTQTC--HPFLTSLKYSFQTTDRICFVME	236						
QY	178	YINGELEYHLQRECFLEPRAREYAAEIASALGYLHNSIVYRDLKPNILDSQGHIV	237						
DB	237	YINGELEYHLQRECFLEPRAREYAAEIASALGYLHNSIVYRDLKPNILDSQGHIV	296						
QY	238	LTDFGLCKENIEHNSSTFGCTGPEYLAPEVHLKQPDYDRTVDMWCLGAVLYEMLYGLPPF	297						
DB	297	ITDFGLCKEDLSYNGTSTFGCTGPEYLAPEVLESDYGRAVDWGTGVVYEMMCGRLPF	356						
QY	298	YSRNTAEYDNLNKLKPNITNSARHLEGLQKDRTKRLGAKDPMFKSHVFFSLINWDDLKKTTP	356						

```

Db 357 YSRDEHVEFLPELILVEEVKFPARLSESARSLLTGLAKDPKRLGGQEDAKEIMEHPFEFA 416
Qy 357 LINWDDLINKITPPNPVNSGNELRHPOPEETEPV- - - - -PNSICKSPDSVLVTASVK 412
Db 417 CINMEDLFNKIEEPKPSKSDTDTRYFDEEFTAEPSVELTPDDNKLSTDSGL- - - - - 470
Qy 413 EAAEAFLGFSYA 424
Db 471 - - - - -PQFEKFSYS 479

RESULT 8
Q24293
ID Q24293 PRELIMINARY; PRT: 530 AA.
AC Q24293;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE DAKT1 serine-threonine protein kinase.
GN AKT1 OR DAKT1 OR CG4006.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94134407; PubMed=8302573;
RA Franke T.F., Tartof K.D., Tschlis P.N.;
RT "The Sh2-like Akt homology (AH) domain of c-akt is present in multiple
RT copies in the genome of vertebrate and invertebrate eucaryotes.
RT Cloning and characterization of the Drosophila melanogaster c-akt
RT homolog Dakt1."
RL Oncogene 9:141-148(1994).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; 226242; CA81204.1; -
DR HSSP; 063450; 1A06.
DR FlyBase; FBgn0010379; Akt1.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 530 AA; 59921 MW; B41E73F2A1833443 CRC64;

Query Match 39.2%; Score 890; DB 5; Length 530;
Best Local Similarity 54.0%; Pred. No. 9.1e-66;
Matches 177; Conservative 46; Mismatches 99; Indels 6; Gaps 4;

Qy 97 DFHLKLVIGKSGFKVLLARHKAEEVFYAVKVLQKAILKKKEKHIMSERVLLKNVXH 156
Db 184 NFEFLKVLGKGTGFKVILCREKATAKLYAIIKILKEVIQKDEVAHTLTESRV-LKSTNH 242
Qy 157 PFLVGLHESFOTADKLYFVLDYINGGELFYHLQRCERCFLEPRARYAAEIASALGYLHSL 216
Db 243 PFLISLKYSFQTNDRCLCFVMQYVNGELFWLHLSHERIFTEDTRFYGAEEIISALGYLHSHQ 302
Qy 217 NIVYRDLPENILLDSQGHIVLTFGLCKENIEHNSTSTFCGTPPEYLAPVLHKKQPYDR 276
Db 303 GIIVYRDLPENILLDSQGHIVLTFGLCKENIEHNSTSTFCGTPPEYLAPVLHKKQPYDR 362

```

```

Qy 277 TVDWMCGLGAVLYEMLYGLPPFYSRNTAEYDNIILNKPLQLKPNITNSARHILLEGILQOKR 336
Db 363 AVDWGSGVVMYEMICSLKFPYKROHDVLTLLILVEEVKFPNITDEAKNLLAGLAKDP 422
Qy 337 TKRL-GAKDDFMEIKSHVFSNLINWDDLINKITPPNPVNSGNELRHPOPEETEPV 395
Db 423 KKRLLGGKDDVKVEIQAHPPFASINWTDVLVKIPPPFPKQVTSDDTKYFKDEFTGHSVE 482
Qy 396 NSICKSPDSVLVTASVKEAAEAFLGFSY 423
Db 483 LT- - - - -PPDTPGPIGSIJAE-EPLFPQFST 506

```

RESULT 9

```

Q8T9A5
ID Q8T9A5 PRELIMINARY; PRT: 530 AA.
AC Q8T9A5;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SDI0374P.
GN AKT1 OR CG4006.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George K.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AY069856; AAL40001.1; -
DR HSSP; P24941; 1BUH.
DR FlyBase; FBgn0010379; Akt1.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 530 AA; 59911 MW; F7508582C7F4A288 CRC64;

```

Query Match

```

Best Local Similarity 39.2%; Score 890; DB 5; Length 530;
Matches 177; Conservative 46; Mismatches 99; Indels 6; Gaps 4;

Qy 97 DFHLKLVIGKSGFKVLLARHKAEEVFYAVKVLQKAILKKKEKHIMSERVLLKNVXH 156
Db 184 NFEFLKVLGKGTGFKVILCREKATAKLYAIIKILKEVIQKDEVAHTLTESRV-LKSTNH 242
Qy 157 PFLVGLHESFOTADKLYFVLDYINGGELFYHLQRCERCFLEPRARYAAEIASALGYLHSL 216
Db 243 PFLISLKYSFQTNDRCLCFVMQYVNGELFWLHLSHERIFTEDTRFYGAEEIISALGYLHSHQ 302
Qy 217 NIVYRDLPENILLDSQGHIVLTFGLCKENIEHNSTSTFCGTPPEYLAPVLHKKQPYDR 276

```


Db 303 GIIYRDLKLENNLLDKGHIKVADEGLCKEDITYGRTRTKTCGTPYLAPEVLDDNDYQ 362
 QY 277 TVDWMLGAVLYEMLYGLPEYSRNTAEYONILANKPLQKLNITNSARHLLGLELOKDR 336
 Db 363 AVDWNGTGVVMYEMICRPLFFNRDHDVLTLLVEEVKFPNITDEAKNLLAGLAKDP 422
 QY 337 TKRL-CAKDDPMEIKSHVFFSLINMDLKNKTKITPPFNPNVSGPNELRHFDPEFTPEVP 395
 Db 423 KRLGGGKDVKEIQAHPEFASINMDLVLLKIPPPKPVQTSDDTRFYDFKFTGESVE 482
 QY 396 NSTGKSPDSVLVTASVKEAAEAFGLFSY 423
 Db 483 LT---PPDPTGPGLSIAE-EPLFPQFSY 506
 RESULT 10
 Q24469 PRELIMINARY: PRT; 611 AA.
 AC Q24469; Q24470: Q9VEY7;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-JUN-2000 (TREMBLrel. 14, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE RAC protein kinase DRAC-PK85.
 GN AKTI OR CG4006.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 RX [1]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 HC STRAIN=BERKELEY;
 MC MEDLINE=20196006; Pubmed=10731132;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazei R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Ball J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Bencos P.V., Berman B.P., Bhandari P., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster A., Gabrielian A.E., Garg N.S., Geibart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Keichum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Lin X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Paila,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RA Andjelkovic M., Jones P.F., Grossniklaus U., Cron P., Schier A.F.,
 RA Dick M., Bilbe G., Hemmings B.A.;
 RL J. Biol. Chem. 0:0-0(1995).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A AND B (SHOWN HERE); MAY BE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 DR EMBL; AE003711; AAF55276.1; -;
 DR EMBL; X83510; CAA58499.2; -;
 DR EMBL; X83510; CAA58500.1; -;
 DR HSSP; Q63450; 1A06.
 DR FlyBase; FBgn010379; Akt1.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00433; pkinase_C; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00220; S-TRC; 1.
 DR SMART; SM00133; S-TK_X; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Alternative splicing; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 FT VARSPIC 1 81 MISSING (IN ISOFORM A).
 SQ SEQUENCE 611 AA; 68514 MW; C1393E43CC27AC34 CRC64;
 Query Match 39.2%; Score 890; DB 5; Length 611;
 Best Local Similarity 54.0%; Pred. No. 1.le-65;
 Matches 177; Conservative 46; Mismatches 99; Indels 6; Gaps 4;
 QY 97 DFHFKVTKGSGFKVLARHKAEEVFYAVKVLQKAILKKKEKHIMSEKNVLLKNVKH 156
 DB 265 NFEEFLKVLGKGTGKVLCKREKAKYAIKILKEVIQDEVAHTLTESRV-LKSTNH 323
 QY 157 PFLVGLHFSFQVAKLYFVLQYINGGELFYHLQRECFLEPRARFYAAEASALGYIHL 216
 DB 324 PFLISLKSFTQNDRLCFYQVYNGELFWLHLSHERIFTEDTKTFYGAELISALGYLH 383
 QY 217 NIVYEDLPENLLDSQGHVLTDFGLCKENIENHSITSTFCGTEYLAFVLUKQYOR 276
 DB 384 GIIYRDLKLENNLLDKGHIKVADEGLCKEDITYGRTRTKTCGTPYLAPEVLDDNDYQ 443
 QY 277 TVDWMLGAVLYEMLYGLPEYSRNTAEYONILANKPLQKLNITNSARHLLGLELOKDR 336
 DB 444 AVDWNGTGVVMYEMICRPLFFNRDHDVLTLLVEEVKFPNITDEAKNLLAGLAKDP 503
 QY 337 TKRL-CAKDDPMEIKSHVFFSLINMDLKNKTKITPPFNPNVSGPNELRHFDPEFTPEVP 395
 DB 504 KRLGGGKDVKEIQAHPEFASINMDLVLLKIPPPKPVQTSDDTRFYDFKFTGESVE 563
 QY 396 NSTGKSPDSVLVTASVKEAAEAFGLFSY 423
 DB 564 LT---PPDPTGPGLSIAE-EPLFPQFSY 587
 RESULT 11
 Q8INB9 PRELIMINARY: PRT; 611 AA.
 AC Q8INB9;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)
 DE CG4006-PC.
 GN AKTI.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

DR PROSITE; P50003; PH_DOMAIN; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TKC_X; 1.
DR PROSITE; P500107; PROTEIN_KINASE_DOM; 1.
DR PROSITE; P500107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; P500108; PROTEIN_KINASE_ST; 1.
DR PROSITE; P500108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; P500108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 479 AA; 55502 MW; AC3121CD9B869F10 CRC64;

Query Match 38.4%; Score 872; DB 13; Length 479;
Best Local Similarity 45.6%; Pred. No. 2.5e-64;
Matches 177; Conservative 68; Mismatches 125; Indels 18; Gaps 6;

QY 39 IQKIANNYSACKHPVQSGILKISQPEPELMANPSPSPSQINILG-PSSNPHAKPSD 97
Db 103 IQAVANG-----LKSREDEPMDINFG-SPGDNLEGMEAAITKSTKVTMSD 149
QY 98 PHELVIGKSGKSVLLARHKAEEVYAVKVLQKKAILKKKEKHIMSERVLLKNVKNHP 157
Db 150 FQYLKLLGKGTGKVLVREKATGMYAMKILREVIILAKDEVAHTITSRV-LQNRHP 208
QY 158 FLVGLHFSQTDADKLYFLVDLNGINGELFYHQRCEFLPRARFYAAEIASALGYHSLN 217
Db 209 FLTLTKYAFQTRDLRCFCVMEYANGCELFHLSRERVFTRDRFYAGAEIVSLYHSD 268
QY 218 IYVRDLKPNILLDSQGHVLDLDFGLCENIEHNSTSTFCGTEYLAPEVLHQPDR 277
Db 269 IYVRDLKLENLMDKDGKIKITDFGLCEGNTNEATMKTFCGTEYLAPEVLHQPDR 328
QY 278 VDWMCGLAVLYEMLYGLPPFSRNTAEYMDNLNKLQKPNITNSARHLEGLQKDR 337
Db 329 VDWMLGLVYVMEYMGCRPLFPYQNDHERLEFELLIMEEIRFPRNLSPKALLGLKKDPK 388
QY 338 KRL-GAKDDPMELKSHVFFSLINWDLINKKTIPTFPNPNVSGPNELRHFDEFTPEVP 396
Db 389 QRLGGGPEDAKEVMTHTKFNNNWQVQLQKLVPPKPKQVTSYDTRFYDEFTQAQITV 448
QY 397 STGKSPDVLVTASVKEAAEAFGLFSYA 424
Db 449 TPDQYDS--LDREDDTRTHFSQFSYS 474

RESULT 13
Q57513 ID O57513 PRELIMINARY; PRT; 480 AA.
AC O57513;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine/threonine protein kinase.
GN AKT1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Theelen M., Swinkels S.U.M., de Jong M.D.M., Thomas A.A.M.,
RA Verkleij A.J., Hanafusa H., Humbel B.;
RT "The differential expression of Akt during the cell cycle is regulated
RT at the translational level."
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF039943; AAB94767.1; -;
DR HSP; P05132; ICTP.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00089; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00133; S_TKC_X; 1.
DR PROSITE; P50003; PH_DOMAIN; 1.
DR PROSITE; P500107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; P500108; PROTEIN_KINASE_ST; 1.
DR PROSITE; P500108; PROTEIN_KINASE_DOM; 1.
DR PROSITE; P500108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 480 AA; 55723 MW; 3B83A47FD6689626 CRC64;

Query Match 38.1%; Score 865; DB 13; Length 480;
Best Local Similarity 43.5%; Pred. No. 9.7e-64;
Matches 181; Conservative 75; Mismatches 136; Indels 22; Gaps 8;

QY 28 MKORRGLNDFIQK-----IANNYSACKHP-----VQSILKISQPEPELMANP 73
Db 63 MTERPKPNTFIIRCLQWTTVIERTFHVETPEREETWKAQTVDLSLKKKEEMDFRS 122
QY 74 SPPP--SPSQINILGPNPNH-AKPSDFHLKVIKIGKSGFGLVLLARHKAEEVYAVKVLQ 130
Db 123 GSPDNGSAEGMEVSTKPKHKVTWNEFEYLLKLGKTFGKVLVKEKATGKYAMKILK 182
QY 131 KKAILLKKKEKHIMSERVLLKNVKNHPFLVGLHFSQTDADKLYFLVDLNGINGELFYHQLR 190
Db 183 KEVIVAKDVAHTLTE-NRVLNSRHPLTALKYISFQTHDRLCFVMEYANGELFFHLSR 241
QY 191 EECFLPRARFYAAEIASALGYLHS-LNIVYRDLKPNILLDSQGHVLDLDFGLCENIEHN 249
Db 242 ERVFSRDRFYAGAEIVSALDYHSEKNVYRDLKLENLMDKDGKIKITDFGLCKEGLK 301
QY 250 HNSTSTFCGTEYLAPEVLHQPDRYVDMWCLGAVLYEMLYGLPPFSRNTAEYMDNLNKL 309
Db 302 DGATMKTFCGTEYLAPEVLHQPDRYVDMWCLGAVLYEMLYGLPPFSRNTAEYMDNLNKL 361
QY 310 LNKPLQKPNITNSARHLEGLQKDRKRL-GAKDDPMELKSHVFFSLINWDLINKKTIPT 368
Db 362 LMEIRFPRTYLSPEAKSLSGLLKKPKORLGGGPDPAKEIMQHREFFAGIVQDVYVYKGL 421
QY 369 TFPFNPVSGPNELRHFDEFTPEVPNPNVSGPDSVLVTASVKEAAEAFGLFSYA 424
Db 422 VDPFKPKQVTSYDTRFYDEFTQAQITVTPPDQDS--MDCVDNKRERHFFPQFSYS 475

RESULT 14
Q98TY9 ID Q98TY9 PRELIMINARY; PRT; 481 AA.
AC Q98TY9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Akt.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersen C.B., Sakae H., Roth R.A., Conti M.;
RT "The protein kinase B/Akt is required for resumption of meiosis in
RT Xenopus oocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF317656; AAG59601.1; -;
DR HSP; P05132; ICTP.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00433; pkinase_C; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 08:14:59 ; Search time 61.0429 Seconds
(without alignments)
298.740 Million cell updates/sec

Title: US-10-000-039A-2

Perfect score: 2270

Sequence: 1 MTVTEAAKGLTYSRMGM.....KEAAEALGFSYAPPTDSFL 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2270	100.0	431	4	US-09-031-295-2
2	2254	99.3	431	2	Sequence 2, Appli
3	2254	99.3	431	3	Sequence 5, Appli
4	2254	99.3	431	3	Sequence 5, Appli
5	2204.5	97.1	430	3	Sequence 9, Appli
6	2204.5	97.1	430	3	Sequence 9, Appli
7	2204.5	97.1	430	3	Sequence 9, Appli
8	850.5	37.5	426	3	Sequence 71, Appli
9	850	37.4	480	3	Sequence 2, Appli
10	850	37.4	727	4	Sequence 139, App
11	806.5	35.5	482	4	Sequence 2, Appli
12	806.5	35.5	495	4	Sequence 3, Appli
13	803	35.4	525	2	Sequence 7, Appli
14	803	35.4	525	2	Sequence 8, Appli
15	803	35.4	525	4	Sequence 16, Appli
16	780.5	34.4	916	4	Sequence 73, Appli
17	780	34.4	637	4	Sequence 2, Appli
18	774.5	34.1	737	4	Sequence 4, Appli
19	773.5	34.1	587	1	Sequence 2, Appli
20	771	34.0	671	6	Patent No. 5266464
21	752.5	33.1	584	4	Sequence 2, Appli
22	697.5	30.7	343	4	Sequence 38, Appli
23	697.5	30.7	595	4	Sequence 69, Appli
24	692.5	30.5	350	4	Sequence 37, Appli
25	692.5	30.5	351	4	Sequence 6, Appli
26	691	30.4	1151	4	Sequence 11, Appli
27	690.5	30.4	576	3	Sequence 2, Appli

28 688.5 30.3 351 4 US-09-394-455-4 Sequence 4, Appli
29 687.5 30.3 336 4 US-09-394-455-2 Sequence 2, Appli
30 687.5 30.3 343 4 US-09-394-455-34 Sequence 34, Appli
31 681.5 30.0 343 4 US-09-394-455-15 Sequence 15, Appli
32 669.5 29.5 264 2 US-07-857-224B-15 Sequence 10, Appli
33 666.5 29.4 264 2 US-07-857-224B-10 Sequence 12, Appli
34 666 29.3 268 4 US-09-430-564-12 Sequence 12, Appli
35 665.5 29.3 264 2 US-07-857-224B-12 Sequence 13, Appli
36 661.5 29.1 263 2 US-07-857-224B-13 Sequence 17, Appli
37 661.5 29.1 264 2 US-07-857-224B-17 Sequence 14, Appli
38 658 29.0 269 2 US-07-857-224B-14 Sequence 11, Appli
39 657.5 29.0 264 2 US-07-857-224B-11 Sequence 16, Appli
40 654 28.8 269 2 US-07-857-224B-16 Sequence 6, Appli
41 649.5 28.6 258 4 US-09-430-564-6 Sequence 4, Appli
42 636.5 28.0 260 2 US-07-857-224B-4 Sequence 3, Appli
43 632.5 27.9 260 2 US-07-857-224B-3 Sequence 2, Appli
44 624.5 27.5 260 2 US-07-857-224B-2 Sequence 8, Appli
45 620.5 27.3 260 2 US-07-857-224B-8

ALIGNMENTS

RESULT 1
US-09-031-295-2
; Sequence 2, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-S3K
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercock, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 058315/0123
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-295-2

Query Match 100.0%; Score 2270; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 4.2e-209;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGINDFIQIANNNSACKHPEVQSILKI 60
|||||
DB 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGINDFIQIANNNSACKHPEVQSILKI 60
|||||

QY 61 SQQPELMNANPPSPSQOINLGPSSNPHAKPSDFHLKVIKGSFGKVLARHKA 120
DB 61 SQQPELMNANPPSPSQOINLGPSSNPHAKPSDFHLKVIKGSFGKVLARHKA 120
QY 121 EVFVAVKYLQKAILKKKEKHINSEKRVNLLKVKHPPFLVGLHFSFQADKLYFVLDYIN 180
DB 121 EVFVAVKYLQKAILKKKEKHINSEKRVNLLKVKHPPFLVGLHFSFQADKLYFVLDYIN 180
QY 181 GGELFYHLQRECFLEPRARYAABIASALGYLSLNIVYRDLPENILLDSQGHVLT 240
DB 181 GGELFYHLQRECFLEPRARYAABIASALGYLSLNIVYRDLPENILLDSQGHVLT 240
QY 241 FGLCKENHSTSTTFCGTPPEYLAPEVHLKQPYDRTVDMWCLGAVLYEMLYGLPPFYSR 300
DB 241 FGLCKENHSTSTTFCGTPPEYLAPEVHLKQPYDRTVDMWCLGAVLYEMLYGLPPFYSR 300
QY 301 NTAEMYDNILNKPQLKPNITNSARHLLLEGLLQKDRTRKLGAKDDFMEIKSHVFFSLINW 360
DB 301 NTAEMYDNILNKPQLKPNITNSARHLLLEGLLQKDRTRKLGAKDDFMEIKSHVFFSLINW 360
QY 361 DDLINKKITPPFNPNVSGPNELRHDFDETEEPVNSICKSPDSVLVTASVKEAAEAF 420
DB 361 DDLINKKITPPFNPNVSGPNELRHDFDETEEPVNSICKSPDSVLVTASVKEAAEAF 420
QY 421 FSYAPPTDSFL 431
DB 421 FSYAPPTDSFL 431

RESULT 2

US-08-712-709-5
; Sequence 5, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-5

Query Match 99.3%; Score 2254; DB 2; Length 431;
Best Local Similarity 99.3%; Pred. No. 1.4e-207;
Matches 428; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MTKTERAAKGLTYSRMKGMVAIIIAFMKQRRMGLNDFIQKIANNSSACKHPEVQSILKI 60
DB 1 MAVTERAAKGLTYSRMKGMVAIIIAFMKQRRMGLNDFIQKIANNSSACKHPEVQSILKI 60
QY 61 SQQPELMNANPPSPSQOINLGPSSNPHAKPSDFHLKVIKGSFGKVLARHKA 120
DB 61 SQQPELMNANPPSPSQOINLGPSSNPHAKPSDFHLKVIKGSFGKVLARHKA 120
QY 121 EVFVAVKYLQKAILKKKEKHINSEKRVNLLKVKHPPFLVGLHFSFQADKLYFVLDYIN 180
DB 121 EVFVAVKYLQKAILKKKEKHINSEKRVNLLKVKHPPFLVGLHFSFQADKLYFVLDYIN 180
QY 181 GGELFYHLQRECFLEPRARYAABIASALGYLSLNIVYRDLPENILLDSQGHVLT 240
DB 181 GGELFYHLQRECFLEPRARYAABIASALGYLSLNIVYRDLPENILLDSQGHVLT 240
QY 241 FGLCKENHSTSTTFCGTPPEYLAPEVHLKQPYDRTVDMWCLGAVLYEMLYGLPPFYSR 300
DB 241 FGLCKENHSTSTTFCGTPPEYLAPEVHLKQPYDRTVDMWCLGAVLYEMLYGLPPFYSR 300
QY 301 NTAEMYDNILNKPQLKPNITNSARHLLLEGLLQKDRTRKLGAKDDFMEIKSHVFFSLINW 360
DB 301 NTAEMYDNILNKPQLKPNITNSARHLLLEGLLQKDRTRKLGAKDDFMEIKSHVFFSLINW 360
QY 361 DDLINKKITPPFNPNVSGPNELRHDFDETEEPVNSICKSPDSVLVTASVKEAAEAF 420
DB 361 DDLINKKITPPFNPNVSGPNELRHDFDETEEPVNSICKSPDSVLVTASVKEAAEAF 420
QY 421 FSYAPPTDSFL 431
DB 421 FSYAPPTDSFL 431

RESULT 3

US-09-111-444-5
; Sequence 5, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,444
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/712,709
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-111-444-5

Query Match 99.3%; Score 2254; DB 3; Length 431;

Best Local Similarity 99.3%; Pred. No. 1.4e-207;

Matches 428; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MTVKTAAGKTLTYSRMGVAILIAFMKQRRGLNDFIQKIANNYSACKHPEVQSILKI 60
Db 1 MAVKTEAAGKTLTYSRMGVAILIAFMKQRRGLNDFIQKIANNYSACKHPEVQSILKI 60
QY 61 SQQPELMANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
Db 61 SQQPELMANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPFLVGLHFSQTADKLYFVLDYIN 180
Db 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPFLVGLHFSQTADKLYFVLDYIN 180
QY 181 GSELFYHLQRCFLPRARFYAAEIASALGYLHLSNIVYRLKPNILLDSQGHVLTLD 240
Db 181 GSELFYHLQRCFLPRARFYAAEIASALGYLHLSNIVYRLKPNILLDSQGHVLTLD 240
QY 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKQPYDRTVDWCLGAVLYEMLYGLPPFYSR 300
Db 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKQPYDRTVDWCLGAVLYEMLYGLPPFYSR 300
QY 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGLOKQRTKRLGAKDDPFMEIKSHVFFSLINW 360
Db 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGLOKQRTKRLGAKDDPFMEIKSHVFFSLINW 360
QY 361 DDLINKKITPPFNPNVSGPNELRHDFPEETEEPPVNSIGKSPDSVLVTASVKEAAEAFILG 420
Db 361 DDLINKKITPPFNPNVSGPNELRHDFPEETEEPPVNSIGKSPDSVLVTASVKEAAEAFILG 420
QY 421 FSYAPPTDSFL 431
Db 421 FSYAPPTDSFL 431
```

RESULT 4

US-09-541-228-5

Sequence 5, Application US/09541228

Patent No. 6232077

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/541,228

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-541-228-5

Query Match 99.3%; Score 2254; DB 3; Length 431;

Best Local Similarity 99.3%; Pred. No. 1.4e-207;

Matches 428; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MTVKTAAGKTLTYSRMGVAILIAFMKQRRGLNDFIQKIANNYSACKHPEVQSILKI 60
Db 1 MAVKTEAAGKTLTYSRMGVAILIAFMKQRRGLNDFIQKIANNYSACKHPEVQSILKI 60
QY 61 SQQPELMANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
Db 61 SQQPELMANPSPSPSQINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPFLVGLHFSQTADKLYFVLDYIN 180
Db 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPFLVGLHFSQTADKLYFVLDYIN 180
QY 181 GSELFYHLQRCFLPRARFYAAEIASALGYLHLSNIVYRLKPNILLDSQGHVLTLD 240
Db 181 GSELFYHLQRCFLPRARFYAAEIASALGYLHLSNIVYRLKPNILLDSQGHVLTLD 240
QY 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKQPYDRTVDWCLGAVLYEMLYGLPPFYSR 300
Db 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKQPYDRTVDWCLGAVLYEMLYGLPPFYSR 300
QY 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGLOKQRTKRLGAKDDPFMEIKSHVFFSLINW 360
Db 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGLOKQRTKRLGAKDDPFMEIKSHVFFSLINW 360
QY 361 DDLINKKITPPFNPNVSGPNELRHDFPEETEEPPVNSIGKSPDSVLVTASVKEAAEAFILG 420
Db 361 DDLINKKITPPFNPNVSGPNELRHDFPEETEEPPVNSIGKSPDSVLVTASVKEAAEAFILG 420
QY 421 FSYAPPTDSFL 431
Db 421 FSYAPPTDSFL 431
```

RESULT 5

US-08-712-709-9

Sequence 9, Application US/08712709

Patent No. 5863780

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/712,709
 FILING DATE: Filed Herewith
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0118 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 430 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 294637
 US-08-712-709-9

Query Match 97.1%; Score 2204.5; DB 2; Length 430;
 Best Local Similarity 96.8%; Pred. No. 8e-203;
 Matches 417; Conservative 8; Mismatches 5; Indels 1; Gaps 1;
 QY 1 MTVKTEAAKGLTYSRMKGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 DB 1 MTVKTEAAKGLTYSRMKGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 QY 61 SQQPELMANPSPSPSPSQIINLGPSSNPHAKPSDFHFLKVIKSGSGFKVLLAKHAE 120
 DB 61 SQQPELMANPSPSPSPSQIINLGPSSNPHAKPSDFHFLKVIKSGSGFKVLLAKHAE 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPPFLVGLHFSFQADKLYFVLDYIN 180
 DB 121 EAFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPPFLVGLHFSFQADKLYFVLDYIN 180
 QY 181 GGELFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILDSQGHIVLTD 240
 DB 181 GGELFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILDSQGHIVLTD 240
 QY 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKOPDYRTVDMWCLGAVLYEMLYGLPPYSR 300
 DB 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKOPDYRTVDMWCLGAVLYEMLYGLPPYSR 300
 QY 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGGLQKORTKRLGAKDDFMEIKSHVFFSLINW 360
 DB 301 NTAEMYDNILNKPLQK - NITNSARHLLLEGGLQKORTKRLGAKDDFMEIKSHVFFSLINW 359
 QY 361 DDLINKKITPPNPVNSGPNELRHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGL 420
 DB 360 DDLINKKITPPNPVNSGPNELRHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGL 419
 QY 421 FSYAPPTDSFL 431
 DB 420 FSYAPPTDSFL 430

RESULT 6

US-09-111-444-9
 ; Sequence 9, Application US/09111444
 ; Patent No. 6045792
 ; GENERAL INFORMATION:

APPLICANT: Au-Young, Janice
 APPLICANT: Guegler, Karl J.
 APPLICANT: Hawkins, Phillip K.
 TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/111,444
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/712,709
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0118 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 430 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 294637
 US-09-111-444-9

Query Match 97.1%; Score 2204.5; DB 3; Length 430;
 Best Local Similarity 96.8%; Pred. No. 8e-203;
 Matches 417; Conservative 8; Mismatches 5; Indels 1; Gaps 1;
 QY 1 MTVKTEAAKGLTYSRMKGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 DB 1 MTVKTEAAKGLTYSRMKGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 QY 61 SQQPELMANPSPSPSPSQIINLGPSSNPHAKPSDFHFLKVIKSGSGFKVLLAKHAE 120
 DB 61 SQQPELMANPSPSPSPSQIINLGPSSNPHAKPSDFHFLKVIKSGSGFKVLLAKHAE 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPPFLVGLHFSFQADKLYFVLDYIN 180
 DB 121 EAFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPPFLVGLHFSFQADKLYFVLDYIN 180
 QY 181 GGELFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILDSQGHIVLTD 240
 DB 181 GGELFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILDSQGHIVLTD 240
 QY 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKOPDYRTVDMWCLGAVLYEMLYGLPPYSR 300
 DB 241 FGLCKENIEHNSTSTFCGTPEYLAPEVLHKKOPDYRTVDMWCLGAVLYEMLYGLPPYSR 300
 QY 301 NTAEMYDNILNKPLQKPNITNSARHLLLEGGLQKORTKRLGAKDDFMEIKSHVFFSLINW 360
 DB 301 NTAEMYDNILNKPLQK - NITNSARHLLLEGGLQKORTKRLGAKDDFMEIKSHVFFSLINW 359
 QY 361 DDLINKKITPPNPVNSGPNELRHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGL 420
 DB 360 DDLINKKITPPNPVNSGPNELRHFDPEFTPEEPVNSIGKSPDSVLVTASVKEAAEAFGL 419

QY 421 FSYAPPTDSFL 431
Db 420 FSYAPPTDSFL 430

RESULT 7
US-09-541-228-9
; Sequence 9, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 294637

US-09-541-228-9

Query Match 97.1%; Score 2204.5; DB 3; Length 430;
Best Local Similarity 96.8%; Pred. No. 8e-203;
Matches 417; Conservative 8; Mismatches 5; Indels 1; Gaps 1;
QY 1 MTVKTEAAAGTITYSMRGWVAILIAFMKQRRMGLNDFIQKLANNYSACKHPEVQSILKI 60
Db 1 MTVKTEAAAGTITYSMRGWVAILIAFMKQRRMGLNDFIQKLANNYSACKHPEVQSILKI 60
QY 61 SQOPEELMANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAKAE 120
Db 61 SQOPEELMANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAKAE 120
QY 121 EYFYAVKVLQKKAALKKKEKHIMSERNVLLKNVHPFLVGLHFSFQADKLYFVLDYIN 180
Db 121 EYFYAVKVLQKKAALKKKEKHIMSERNVLLKNVHPFLVGLHFSFQADKLYFVLDYIN 180
QY 181 GGELFYHLQRCRCFLEPRARFYAAETASALGYLHSLNIYVYRDLPENILLDSQGHIVLTD 240
Db 181 GGELFYHLQRCRCFLEPRARFYAAETASALGYLHSLNIYVYRDLPENILLDSQGHIVLTD 240

QY 241 FGLCKENIEHNSTSTFCGTPEYLAPVHLHKOPYDRTVDWCLGAVLYEMLYGLPPFYSK 300
Db 241 FGLCKENIEHNSTSTFCGTPEYLAPVHLHKOPYDRTVDWCLGAVLYEMLYGLPPFYSK 300
QY 301 NTAEMYDNLNKPLOLKPNI TNSARHLLEGLLQKORTKRLGAKDDPMEIKSHVFFSLINW 460
Db 301 NTAEMYDNLNKPLOLKPNI TNSARHLLEGLLQKORTKRLGAKDDPMEIKSHVFFSLINW 359
QY 361 DDLINKKITPPNPVNSGPNELRHDFPEETPEVPVNSICKSPDVLVTFASVKAAPAFIAG 420
Db 360 DDLINKKITPPNPVNSGPNELRHDFPEETPEVPVNSICKSPDVLVTFASVKAAPAFIAG 419
QY 421 FSYAPPTDSFL 431
Db 420 FSYAPPTDSFL 430

RESULT 8

US-09-417-197-71
; Sequence 71, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole THASTRUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To A
; FILE OF INVENTION: On A Cellular Response
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417,197
; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 71
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PKB-EGFP fusion
US-09-417-197-71

Query Match 37.5%; Score 850.5; DB 4; Length 726;
Best Local Similarity 42.8%; Pred. No. 9e-73;
Matches 185; Conservative 72; Mismatches 132; Indels 43; Gaps 10;
QY 28 MKORRMGLNDFIOLK-----IANNYSACKHPE-----VQSILKISQPEPELIMANP 73
Db 63 MKTERPRPNTFTIRCLQWTVIETKTHVETPEEREGWTTAQTVAIDGLKKQEEEMDPHS 122
QY 74 SPPSPS-----QQINLGPSSNPHAKP-----SDHFLLKVIKGSFGKVLARHAKAEVFP 123
Db 123 GSPSDNSGAEMEVSLE-----AKPKHRTVTHNEFEYLLKLGKGTGFKVILVKEKATGRY 175
QY 124 YAVKVLQKKAALKKKEKHIMSERNVLLKNVHPFLVGLHFSFQADKLYFVLDYINNGE 183
Db 176 YAMKILKKEVIVAKUEVAHTLIE-NRVLQNSRHPFLTALKYSFQTHDRLCFVMEYANGGE 234
QY 184 LEYHLQRCRCFLEPRARFYAAETASALGYLHSLNIYVYRDLPENILLDSQGHIVLTD 242
Db 235 LEFHLSRERFVSEDARFYGAIEVSALDYHSEKNVYRDLPENILLDSQGHIVLTD 294
QY 243 LCKENIEHNSTSTFCGTPEYLAPVHLHKOPYDRTVDWCLGAVLYEMLYGLPPFYSKNT 302
Db 295 LCKEGIKGATGATKTKTCGTPEYLAPVHLNEDNDYGRAVDWGLGVNVMYEMMCGRLPFPYNDH 354
QY 303 ARMYONITLKNPQLKPNITNSARHLLEGLLQKORTKRLGAKDDPMEIKSHVFFSLINWD 361
Db 355 EKLFEILLMEELRFPRTLGPEAKSLLSGLLKKDKPQKRGGSSEDAKEIMQHRFFAGIVNQ 414
QY 362 DLINKKITPPNPVNSGPNELRHDFPEETPEVPVNSICKSPDVLVTFASVKAAPAFIAG 421
Db 415 HVYKELSPPEKPOVTSETDTRYFDEETQAQMITTPPDQDDSMCEVDS--ERRPHFPQF 472
QY 422 SYA-----PP 426
Db 473 SYSASSTASDPP 484

RESULT 9

```

US-09-091-058-2
; Sequence 2, Application US/09091058
; Patent No. 6054285
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Frech, Matthias
; TITLE OF INVENTION: Screening Method
; FILE REFERENCE: 4-20683/A/20684/PCT
; CURRENT APPLICATION NUMBER: US/09/091,058
; CURRENT FILING DATE: 1998-06-10
; EARLIER APPLICATION NUMBER: PCT/EP96/04814
; EARLIER FILING DATE: 1996-11-05
; EARLIER APPLICATION NUMBER: 9525703.6
; EARLIER FILING DATE: 1995-12-15
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-091-058-2

```

Query Match 37.48; Score 850; DB 3; Length 480;

```

best local similarity 43.36; 72; Mismatches 132; Indels 36; Gaps 9;
Matches 183; Conservative
Qy 28 MKORRMGLNDPIOK-----IANN5YACKHPE-----VQSILKISQPOPELMMANP 73
Db 11 11 11 : : : 11 : : : : : : : : : : : : : : : :
63 KTERPRPNTFIIRCLQWTVVIENTPHVETPEEREEMTTAQTADVADGLKKOEEMEDFRS 122
Qy 74 SPSPSS---QQINLGCPSSNPHAKP-----SDFHFLKVIKGSFGKVLARHKAEBVF 123
Db 11 : : : : : : : : : : : : : : : : : : : : : : : : : :
123 GSPSDMSGABEEMEVSL-----AKPKHRTVMNEFYELKLLGKOTFGKVLAVKKGATGRY 175
Qy 124 YAVVLQKKAKLKKKEEKHIMSENRVLLKNVKKHPFLVGLHFSFOTADKLXVLDYINGGE 183
Db 11 11 11 : : : : : : : : : : : : : : : : : : : : : :
176 YAMKILKKKEIVAKDEVAHTLTE-NRVLQNSRHPFLTALKVSFOTHORLCFVMEYANGGE 234
Qy 184 LFVHLQRECFLEPRARFYAAETASALGYLHS-LNIVYRDLKPNILLDSOGHIVLTDG 242
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
235 LFFHLSRERVFSEDRARFYGAETVSALDYLHSEKNVYVRDLKLENLMDKDGHTKIITDVG 294
Qy 243 LCKENIBHNSTTTECGTPTVLAPEVLHKKOPYDRTVDWMCGLGAVLYEMLYGLPFPYSRNT 302
Db 11 11 : : 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
295 LCHEGIDGATMKTFCGTPTVLAPEVLEDNDYGRAVDWGLGVVYEMCGRLFPFYMDH 354
Qy 303 AEWYDNLNKLPLQKPNITNSARHLLBGLQKQKTKRL-CAKQDFMEIKSHVFFSLINWD 361
Db 11 : : : : : : : : : : : : : : : : : : : : : : : : : :
355 EKLFELTMBEIRFPPTLGPPEAKSLLSGLLKKPKQKRLGGGSEDAKEIMQHRFAGIVWQ 414
Qy 362 DLNKKIKITPPENNVGSPNELRHFDPTEEPVPSNGKSPDSVLYVTASVYKAAEAFLGF 421
Db 11 11 11 11 : : : : : : : : : : : : : : : : : : : : : :
415 HVYEKLSPPFKQVTSSETDTRVDFDEFTAQMITITPPDOODSMCECVDS--ERRPHFPQF 472
Qy 422 SYA 424
Db 11 :
473 SYS 475

```

RESIST 10

```

US-09-417-197-139
: Sequence 139, Application US/09417197
: Patent No. 6518021
: GENERAL INFORMATION:
: APPLICANT: Ole THASTRUP, et al.
: TITLE OF INVENTION: A Method For Ext
: TITLE OF INVENTION: On A Cellular R
: FILE REFERENCE: 3759-0110P
: CURRENT APPLICATION NUMBER: US/09/41
: CURRENT FILING DATE: 1999-10-07

```

```

; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 139
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PKB f
US-09-417-197-139

```

Query Match 37.4%; Score 850; DB 4; Length 727;
Best Local Similarity 43.3%; Pred. No. 1e-72;
Matches 183; Conservative 72; Mismatches 132; Indels 36; Gaps 9;

Qy	28	MKQRMGLNDFQK-----TANNSYACKHPE-----VQSILKSQOPELNMNP	73
Db	310	MTTERPRNPFTIIRCLQWTFIERTFHVETPEEREEMTTAIQTVAOGLKQBEEDNFRS	369
Qy	74	SPSPS---QOINLGSSNPHAKP-----SDFHELKVIKGSCKVLIAHKAFFV	123
Db	370	GSFSDNSGAEBEVS-----AKPHRYTMNFEYKLGKLGKGTGKVLIVKEKATGY	422
Qy	124	YAVVLORKAIKKKEEKHINSERNVLLKNWKHPFLVGLHFSEQTADKLYFLVDYINGSE	183
Db	423	YAMKILKREVIIVAKDEVAHTLTE-NKVLQNSRHPPFTALKYSFQTHDRLCFYMEYANGCE	481
Qy	184	LPYHLORERCFLPEPARFYAAEIASALGYLIIS-UNIVYROLKPEINLLDSQGHIVLDFG	242
Db	482	LFHLSRKFVSEDRARYGAEIVSALDYHSEKNVYVYROLKLENLLDKDGHKIKITDFG	541
Qy	243	LCKENIEHNSTSTFCGTPEYLAPVLHKKOPYDRTVDWMCLGAVLYEMLYGPPFYSRMT	302
Db	542	LCKEGIKOGATMKTFCGTPEYLAPVLEDNDYGRAVDWGLGVVYVEMCGRLPFYVQDH	601
Qy	303	AEWONILNKPLQKNPNTNSARHLLBGLQKDKRKL-CAKDDFMEIKSHVFFSLINMD	361
Db	602	EXLFELLIMEEIRFPRTLGPEAKSLLSGLLKDKPKQLGGSEDAKEIMQHRFEAGIVMO	661
Qy	362	DLINKKITPPNPVNSGPNELRHDPFTEEPVNSICKSPDSVLYTASVKEAANEAFIAC	421
Db	662	HYEKKLSPPFAPQVTSSTDTRYFDEEFTAQMITPTPOODDSMECVDS--EKKPHFPGF	719
Qy	422	SYA 424	
Db	720	SVS 722	

```

RESULT 11
US-09-430-564-2
; Sequence 2, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Kay K. Lee-Pruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
; TITLE OF INVENTION: PRIMERS, PROBES, AND DETECTION METHODS,
; FILE REFERENCE: 00246/506002
; CURRENT APPLICATION NUMBER: US/09/430,564
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-564-2

```

Query Match	35.5%	Score	806.5;	DB	4;	Length	482;
Best Local Similarity	44.3%;	Pred. No.	8.1e-69;				
Matches	172;	Conservative	66;	Mismatches	111;	Indels	39;
						Gaps	12;

```

; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
;
; TITLE OF INVENTION: On A Cellular Response
;
; FILE REFERENCE: 3759-0110P
;
; CURRENT APPLICATION NUMBER: US/09/417,197
;
; CURRENT FILING DATE: 1999-10-07

```

QY 65 EPELMANPSP-----PPSPSQ-----INLGPSS-NPHAKPSDFHLKVIK 106
Db 20 EPELSPADACPLAELRAAGLEPVGHYEVELTETSVNNGPERIGPHC-----FELLRLVGK 75
QY 107 GSPGKVLARHKAEEV-----FYAVKVLQKAILAK-KEEKHIMSERVLLKNVHPFLVG 161
Db 76 GYGKVFQVR-KVQGTNLGKIYAMKVLKAKIVRNAKDTAHTRAERNI-LESVKHPFIVE 133
QY 162 LHFSQADKLYFVLDYINGGELFYHLQRECFLEPRARFYAAETASALGYLHLSNIVYR 221
Db 134 LAYAFQGGKLYLILECUGGELFTHLREGIFLEDYACFYLAETLALGHLSQGIYR 193
QY 222 DLKPNILDSQGHVLTDFGLCKENIEHNSSTTFCTGPEVLAPEVLHKOPYDRTVDWM 281
Db 194 DLKPNIMLSQGHVLTDFGLCKESIEHGAHTFTFCGTIEYMAPEILLVRSNGHNAVDMW 253
QY 282 CLGAVLYEMLYGLPPFYSRNTAEYDNLNKPLOLKPNTITNSARHLLGLLQKDKTKKL- 340
Db 254 SLGALMYDLTSGPPTAENRKKTKIIRGKLALPPYLTDPDARDLVKKFLKRNPSQRIG 313
QY 341 GAKDDFMEIKSHVFFSLINWDDLKNTKTPPFPNVSGPNELRHFDPFTFE-PPVNSIG 399
Db 314 GGPDAADVQRHPPFRHNMWDDLAWRVDPFRPCLQSEEDVSQDFTRFTQTPV----- 368
QY 400 KSPDSVLVTASVKEAAEAFGLFSYAPPT 427
Db 369 DSPDD---TALSANOAFGLFTYVAPS 393

RESULT 12
US-09-430-564-3
; Sequence 3, Application US/09430564
; Patent No. 6372467
; GENERAL INFORMATION:
; APPLICANT: John Blenis
; APPLICANT: Key K. Lee-Fruman
; APPLICANT: Calvin J. Kuo
; TITLE OF INVENTION: P54S6K AND P85S6K GENES, PROTEINS,
; FILE REFERENCE: 002467/306002
; CURRENT APPLICATION NUMBER: US/09/430, 564
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/106,141
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 3
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-430-564-3

Query Match 35.58; Score 806.5; DB 4; Length 495;
Best Local Similarity 44.38; Pred. No. 8.5e-69;
Matches 172; Conservative 66; Mismatches 111; Indels 39; Gaps 12;
QY 65 EPELMANPSP-----PPSPSQ-----INLGPSS-NPHAKPSDFHLKVIK 106
Db 33 EPELSPADACPLAELRAAGLEPVGHYEVELTETSVNNGPERIGPHC-----FELLRLVGK 88
QY 107 GSPGKVLARHKAEEV-----FYAVKVLQKAILAK-KEEKHIMSERVLLKNVHPFLVG 161
Db 89 GYGKVFQVR-KVQGTNLGKIYAMKVLKAKIVRNAKDTAHTRAERNI-LESVKHPFIVE 146
QY 162 LHFSQADKLYFVLDYINGGELFYHLQRECFLEPRARFYAAETASALGYLHLSNIVYR 221
Db 147 LAYAFQGGKLYLILECUGGELFTHLREGIFLEDYACFYLAETLALGHLSQGIYR 206
QY 222 DLKPNILDSQGHVLTDFGLCKENIEHNSSTTFCTGPEVLAPEVLHKOPYDRTVDWM 281
Db 207 DLKPNIMLSQGHVLTDFGLCKESIEHGAHTFTFCGTIEYMAPEILLVRSNGHNAVDMW 266

QY 282 CLGAVLYEMLYGLPPFYSRNTAEYDNLNKPLOLKPNTITNSARHLLGLLQKDKTKKL- 340
Db 267 SLGALMYDLTSGPPTAENRKKTKIIRGKLALPPYLTDPDARDLVKKFLKRNPSQRIG 326
QY 341 GAKDDFMEIKSHVFFSLINWDDLKNTKTPPFPNVSGPNELRHFDPFTFE-PPVNSIG 399
Db 327 GGPDAADVQRHPPFRHNMWDDLAWRVDPFRPCLQSEEDVSQDFTRFTQTPV----- 381
QY 400 KSPDSVLVTASVKEAAEAFGLFSYAPPT 427
Db 382 DSPDD---TALSANOAFGLFTYVAPS 406
RESULT 13
US-08-749-902-7
; Sequence 7, Application US/08749902
; Patent No. 5985635
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,902
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0150 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 525 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 189508
US-08-749-902-7

Query Match 35.44; Score 803; DB 2; Length 525;
Best Local Similarity 46.38; Pred. No. 2e-68;
Matches 163; Conservative 64; Mismatches 105; Indels 20; Gaps 9;
QY 83 INLGPSSNPHAKPSDFHLKVIKVGSGKVLARH-----KAEVYFAVKVLQKAILK- 137
Db 79 VNRGPEK---IRPECFELLKVLGKGGVQKVFQYKRVGTANTGKIF-AMKVLKAMIVRNA 134
QY 138 KEEKHIMSERVLLKNVHPFLVGIFHSFQTAOKLYFVLDYINGGELFYHLQRECFLEP 197
Db 135 KOTAHTKAERNI-LEEVKHPFIVULIYAFQGTGKGLYLILEYLSOGELFMQLREGIFMED 193
QY 198 RAEYAAETASALGYLHLSNIVYRDLKPNILDSQGHVLTDFGLCKENIEHNSSTTF 257

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 11, 2003, 06:08:14 ; Search time 136.884 Seconds
(without alignments)
499.774 Million cell updates/sec

Title: US-10-000-039a-2
Perfect score: 2270
Sequence: 1 MPVKTEAAKGLTLYSRMRGM.....KEAAEALGFSVAPTDLSFL 431

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2270	100.0	431	19 AAW77217	Human cell-volume
2	2270	100.0	431	21 AAB24115	Human serum and gl
3	2270	100.0	431	22 AAB65613	Novel protein kina
4	2267	99.9	431	20 AAW90139	Human sgk protein.
5	2261	99.6	431	21 AAY95279	Human serum and gl
6	2254	99.3	431	19 AAW54025	Human protein kina
7	2222	97.9	431	21 AAB24116	Rat serum and gluc
8	2222	97.9	431	21 AAY93530	A rat serum and gl
9	2150.5	94.7	430	22 AAB65614	Novel protein kina

10	1988	87.6	388	20 AAY48573	Human breast tumou
11	1979	87.2	373	22 AAB99815	AGC protein kinase
12	1695	74.7	321	22 AAB99836	AGC protein kinase
13	1477	65.1	496	20 AAW87772	Human serum glucoc
14	1472	64.8	429	21 AAY95278	Human serum and gl
15	1472	64.8	496	22 AAY87228	Novel central nerv
16	1472	64.8	496	22 AAY87382	Novel central nerv
17	1472	64.8	496	22 AAM40620	Human polypeptide
18	1459	64.3	429	21 AAB24114	Human serine threo
19	1457	64.2	496	22 AAB38834	Human polypeptide
20	1415	62.3	427	21 AAY95276	Human serum and gl
21	1409.5	62.1	382	22 AAM25594	Human protein sequ
22	1407	62.0	367	21 AAY95275	Human serum and gl
23	1407	62.0	367	22 AAY28087	Novel human secret
24	1407	62.0	367	23 AAE22765	Novel protein kina
25	1405	61.9	367	22 AAB65708	Mouse serum and gl
26	1383	60.9	367	21 AAY95277	Human polypeptide,
27	1277	56.3	464	22 AAM93732	AGC protein kinase
28	1263.5	55.7	319	22 AAB99838	AGC protein kinase
29	1250.5	55.1	322	22 AAB99837	Novel protein kina
30	1223	53.9	446	22 AAB65616	Novel central nerv
31	964.5	42.5	276	22 AAB97658	Novel signal trans
32	964.5	42.5	276	22 AAB17104	Novel protein kina
33	951.5	41.9	244	22 AAB65615	Amino acid sequenc
34	909	40.0	479	21 AAB19284	Human Akt-3. Homo
35	909	40.0	479	21 AAB13393	Human protein kina
36	909	40.0	479	23 AAB79420	Human Akt3 protein
37	909	40.0	479	23 ABB06998	Amino acid sequenc
38	905	39.9	465	21 AAB19011	Human Akt3. Homo
39	905	39.9	465	22 AAG78018	Human serine/threo
40	905	39.9	465	22 AAB19996	Human serine/threo
41	905	39.9	465	22 AAB19998	AKT3 related polyp
42	903	39.8	454	22 AAG78021	Human OKF2896
43	901	39.7	479	21 AAB43132	Drosophila melanog
44	890	39.2	530	22 ABB66370	Drosophila melanog
45	890	39.2	530	22 ABB71272	Drosophila melanog

ALIGNMENTS

RESULT 1
AAW77217
ID AAW77217 standard; Protein: 431 AA.
AC
AAW77217:
XX
XX
DT 16-NOV-1998 (first entry)
XX
DE Human cell-volume regulating kinase h-sgk.
XX
KW Human; cell-volume; kinase; h-sgk; diabetes mellitus;
KW renal insufficiency; inflammation; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN EP861896-A2.
XX
PD 02-SEP-1998.
XX
PF 27-JAN-1998; 98EP-0101338.
XX
PR 28-FEB-1997; 97DE-1008173.
XX
XX (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI Lang F, Waldegger S;
XX
DR WPI: 1998-449109/39.
DR N-PSDB: AAV48311.
XX
PT New nucleic acid encoding cell-volume regulating kinase h-sgk and related proteins - used for diagnosis and treatment of diseases

100

PT involving changes in cell volume, e.g. renal insufficiency,
 PT inflammation, infections etc.

XX Disclosure: Fig 1; 15pp; German.

XX The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments,
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.

XX Sequence 431 AA;

Query Match 100.0%; Score 2270; DB 19; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.8e-204;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 DB 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60
 QY 61 SQOPEPELMANANPPPPSPSQOINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120
 DB 61 SQOPEPELMANANPPPPSPSQOINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFVLDYIN 180
 DB 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFVLDYIN 180
 QY 181 GGELFYHLQRCFCLEPRAREYAAEIASALGYLHSLNIVYRDLPENILDSQGHVLT 240
 DB 181 GGELFYHLQRCFCLEPRAREYAAEIASALGYLHSLNIVYRDLPENILDSQGHVLT 240
 QY 241 FGLCKENIEHNSTTTCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPFY 300
 DB 241 FGLCKENIEHNSTTTCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPFY 300
 QY 301 NTAEMYDNILNKPLOLKNITNSARHLLGLLQKDRKRLGAKODFMEIKSHVFFSLINW 360
 DB 301 NTAEMYDNILNKPLOLKNITNSARHLLGLLQKDRKRLGAKODFMEIKSHVFFSLINW 360
 QY 361 DDLINKKITPPPNVSGPNELRHDPDTEEPVPNSIGKSPDSVLTASVKEAAEAF 420
 DB 361 DDLINKKITPPPNVSGPNELRHDPDTEEPVPNSIGKSPDSVLTASVKEAAEAF 420
 QY 421 FSYAPPTDSFL 431
 DB 421 FSYAPPTDSFL 431

RESULT 2

AAB24115
 ID AAB24115 standard; Protein; 431 AA.

XX AAB24115;

XX 29-JAN-2001 (first entry)

DE Human serum and glucocorticoid regulated kinase protein (HSGK).

XX Human; STK3; serine threonine kinase; HSGK; RSCK;

KW serum and glucocorticoid regulated kinase.

XX Homo sapiens.

OS CN1259573-A.

XX 12-JUL-2000.

PD

XX 29-OCT-1998; 98CN-0123822.
 XX 29-OCT-1998; 98CN-0123822.
 PR (UYFU-) UNIV FUDAN.

XX Yu L, Fu Q, Zhao Y;

XX WPI; 2000-587991/56.

XX New human serine threonine protein kinase, its code sequence,
 PT preparation, and use -

XX Example 3; Fig 3; 29pp; Chinese.

XX The present invention describes human serine threonine kinase STK3,
 CC which is a new member of the serine/threonine protein kinase family.
 CC Human STK3 shares homology with human serum and glucocorticoid
 CC regulated kinase (HSGK) and rat SGK (RSCK). The present sequence
 CC represents the human HSGK protein sequence from the present invention.

SQ Sequence 431 AA;

Query Match 100.0%; Score 2270; DB 21; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.8e-204;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60

DB 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRMGLNDFIQKIANNACKHPEVQSILKI 60

QY 61 SQOPEPELMANANPPPPSPSQOINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120

DB 61 SQOPEPELMANANPPPPSPSQOINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120

QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFVLDYIN 180

DB 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFVLDYIN 180

QY 181 GGELFYHLQRCFCLEPRAREYAAEIASALGYLHSLNIVYRDLPENILDSQGHVLT 240

DB 181 GGELFYHLQRCFCLEPRAREYAAEIASALGYLHSLNIVYRDLPENILDSQGHVLT 240

QY 241 FGLCKENIEHNSTTTCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPFY 300

DB 241 FGLCKENIEHNSTTTCGTPPEYLAPEVLHKKQPYDRTVDMWCLGAVLYEMLYGLPPFY 300

QY 301 NTAEMYDNILNKPLOLKNITNSARHLLGLLQKDRKRLGAKODFMEIKSHVFFSLINW 360

DB 301 NTAEMYDNILNKPLOLKNITNSARHLLGLLQKDRKRLGAKODFMEIKSHVFFSLINW 360

QY 361 DDLINKKITPPPNVSGPNELRHDPDTEEPVPNSIGKSPDSVLTASVKEAAEAF 420

DB 361 DDLINKKITPPPNVSGPNELRHDPDTEEPVPNSIGKSPDSVLTASVKEAAEAF 420

QY 421 FSYAPPTDSFL 431

DB 421 FSYAPPTDSFL 431

RESULT 3

AAB65613

ID AAB65613 standard; Protein; 431 AA.

XX AAB65613;

XX 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 139.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;

XX

KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Homo sapiens.
 XX WO200073469-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US14842.
 XX 28-MAY-1999; 99US-0136503.
 XX (SUG-) SUGEN INC.
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 DR WPI; 2001-032161/04.
 DR N-PSDB; AAF44639.
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 PS Claim 10; Fig 1; 310pp; English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 XX Sequence 431 AA;

Query Match 100.0%; Score 2270; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.8e-204;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVKTEAAKGTLYSRMRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60
 DB 1 MTVKTEAAKGTLYSRMRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60

QY 61 SQQPELMNANPPSPSQQINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120
 DB 61 SQQPELMNANPPSPSQQINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120

QY 121 EVFYAVKYLQKAILKKKEEKIMSERNVLLKNVHPFLVGLHFSQFADKLYFVLDYIN 180
 DB 121 EVFYAVKYLQKAILKKKEEKIMSERNVLLKNVHPFLVGLHFSQFADKLYFVLDYIN 180

QY 181 GGELFVHLQRECFLEPRARFYAAETASALGYLSINIVYRLKPKENTILDSQGHVIFD 240
 DB 181 GGELFVHLQRECFLEPRARFYAAETASALGYLSINIVYRLKPKENTILDSQGHVIFD 240

QY 241 FGLCKENIEHNSTTFTFCGTPEYLAPEVLHQPYDRTVDWVCLGAVLYEMLYGLPPFYSR 300
 DB 241 FGLCKENIEHNSTTFTFCGTPEYLAPEVLHQPYDRTVDWVCLGAVLYEMLYGLPPFYSR 300

QY 301 NTAEMYDNLNKPQLQKPNITNSARHLLGLEGLQKORTKRLGAKDMEIKSHVFFSLINW 360

DB 301 NTAEMYDNLNKPQLQKPNITNSARHLLGLEGLQKORTKRLGAKDMEIKSHVFFSLINW 360
 QY 361 DDLINKKITPPNPNVSGPNEI.RHFDPETPEPVPNSICKSPDSVLTASVKEAAFAFI 420
 DB 361 DDLINKKITPPNPNVSGPNEI.RHFDPETPEPVPNSICKSPDSVLTASVKEAAFAFI 420
 QY 421 FSYAPPTDSFL 431
 DB 421 FSYAPPTDSFL 431

RESULT 4

AAW90139

ID AAW90139 standard; Protein: 431 AA.

XX

AC AAW90139;

XX

DT 15-MAR-1999 (first entry)

XX

DE Human sgk protein.

XX

KW Serum glucocorticoid regulated kinase; sgk; human; treatment; inhibitor;
 KW serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 KW chronic renal failure; inflammation; Alzheimers disease; wound.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Misc-difference 381

FT /note= "encoded by GAG"

PN EP887081-A2.

XX

XX 30-DEC-1998.

XX

PF 27-MAY-1998; 98EP-0304189.

XX

PR 27-JUN-1997; 97US-0051124.

XX

XX (SMTK) SMITHKLINE BEECHAM CORP.

XX

Kumar JM;

XX

WPI; 1999-047627/05.

XX

DR N-PSDB; AAW74190.

XX

PT Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 PT disease - by administration of nucleic acids and antagonists which
 PT inhibit activity or expression of human serum glucocorticoid
 PT regulated kinase (sgk), a serine/threonine protein kinase

XX

PS Disclosure; Page 16-17; 17pp; English.

XX

CC This sequence represents a novel human serum glucocorticoid regulated
 CC kinase (sgk) protein which is a member of the serine/threonine protein
 CC kinase family. This protein is used for the treatment of a subject having
 CC need to inhibit/antagonise activity or expression of human sgk
 CC polypeptide e.g. for the treatment of chronic renal failure, diabetic
 CC nephropathy, inflammation, Alzheimers disease and wounds.

XX

SQ Sequence 431 AA;

Query Match

Best Local Similarity 99.9%; Score 2267; DB 20; Length 431;

Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTVKTEAAKGTLYSRMRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60

DB 1 MTVKTEAAKGTLYSRMRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60

QY 61 SQQPELMNANPPSPSQQINLGPSNPHAKPSDFHFLKVIKGSFGKVLARHKA 120

Db 61 SQQPELMNANPPSPSQQINLGPSSNPHAKPSDFHFLKVLGKSGFKVLLARHAE 120
 Qy 121 EVFYAVKVLQKKAILKKKEKHIMSERNVLLKNVKKHPFLVGLHFSFQTADKLYFVLDYIN 180
 Db 121 EVFYAVKVLQKKAILKKKEKHIMSERNVLLKNVKKHPFLVGLHFSFQTADKLYFVLDYIN 180
 Qy 181 GGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLKPKENILDSQGHVILTD 240
 Db 181 GGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLKPKENILDSQGHVILTD 240
 Qy 241 FGLCKENIHNSTSTFCGTPPEYLAPEVLHKKQPYDRTVDWMCGLGAVLYEMLYGLPPYSR 300
 Db 241 FGLCKENIHNSTSTFCGTPPEYLAPEVLHKKQPYDRTVDWMCGLGAVLYEMLYGLPPYSR 300
 Qy 301 NTAEMYDNILNKPLQKPNITNSARHLEGLLQKDRTKRLGAKDDFMEIKSHVFFSLINW 360
 Db 301 NTAEMYDNILNKPLQKPNITNSARHLEGLLQKDRTKRLGAKDDFMEIKSHVFFSLINW 360
 Qy 361 DDLINKKITPPFPNVSPGPNELRHFDPEFTEEPVNSIGKSPDSVLTASYKAAEAFLG 420
 Db 361 DDLINKKITPPFPNVSPGPNELRHFDPEFTEEPVNSIGKSPDSVLTASYKAAEAFLG 420
 Qy 421 FSYAPPTDSFL 431
 Db 421 FSYAPPTDSFL 431

RESULT 5

AAW5279
 ID AAW5279 standard; Protein: 431 AA.

XX AC AAW5279;

XX DT 12-SEP-2000 (first entry)

XX DE Human serum and glucocorticoid-induced protein kinase.

XX KW Serum and glucocorticoid-induced protein kinase; SGK; human; phosphorylation; cancer; diabetes; ischaemia; therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 256 /note= "O-phosphorylated"

FT Modified-site 422 /note= "O-phosphorylated"

XX WO200035946-A1.

XX PD 22-JUN-2000.

XX PF 14-DEC-1999; 99WO-GB04232.

XX PR 14-DEC-1998; 98US-0112217.

XX PR 19-AUG-1999; 99GB-0019676.

XX PA (UYDU-) UNIV DUNDEE.

XX PI Cohen P, Kobayashi T, Deak M;

XX WPI; 2000-442364/38.

XX PT Modulation of serum and glucocorticoid-induced protein kinase (SGK) by phosphorylation with 3-phosphoinositide-dependent protein kinase-1 (PK1) or dephosphorylation, useful for treatment of cancer, diabetes and ischemic diseases.

XX PS Disclosure; Fig 13; 127pp; English.

XX CC The present sequence is that of human serum and

CC glucocorticoid-induced protein kinase (SGK). SGK was initially identified as a glucocorticoid and osmotic stress-responsive gene.

CC It is activated by phosphorylation of its Thr-308 and Ser-472 residues. Novel isoforms of SGK, i.e. SGK2 and SGK3 (see AAY9275-78), have been identified. The invention provides methods of activating SGK activity by phosphorylation using 3-phosphoinositide-dependent protein kinase-1 (PK1), and of reducing the activity of SGK by dephosphorylation. The invention also provides a method of identifying a compound that modulates the activity of SGK. Such compounds are useful for treating patients requiring modulation of SGK, such as patients with cancer, diabetes or ischaemic disease.

XX SQ Sequence 431 AA;

Query Match 99.6%; Score 2261; DB 21; Length 431;
 Best Local Similarity 99.8%; Pred. No. 1.9e-203;
 Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTVKTEAAKGLTYSRMRGMVAIIAFMKORRGLNDFIQKIANNYSACKHPEVQSILKI 60

Db 1 MTVKTEAAKGLTYSRMRGMVAIIAFMKORRGLNDFIQKIANNYSACKHPEVQSILKI 60

Qy 61 SQQPELMNANPPSPSQQINLGPSSNPHAKPSDFHFLKVLGKSGFKVLLARHAE 120

Db 61 SQQPELMNANPPSPSQQINLGPSSNPHAKPSDFHFLKVLGKSGFKVLLARHAE 120

Qy 121 EVFYAVKVLQKKAILKKKEKHIMSERNVLLKNVKKHPFLVGLHFSFQTADKLYFVLDYIN 180

Db 121 EVFYAVKVLQKKAILKKKEKHIMSERNVLLKNVKKHPFLVGLHFSFQTADKLYFVLDYIN 180

Qy 181 GGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLKPKENILDSQGHVILTD 240

Db 181 GGELFYHLQRCERCFLEPRARFYAAEIASALGYLHSLNIVYRDLKPKENILDSQGHVILTD 240

Qy 241 FGLCKENIHNSTSTFCGTPPEYLAPEVLHKKQPYDRTVDWMCGLGAVLYEMLYGLPPYSR 300

Db 241 FGLCKENIHNSTSTFCGTPPEYLAPEVLHKKQPYDRTVDWMCGLGAVLYEMLYGLPPYSR 300

Qy 301 NTAEMYDNILNKPLQKPNITNSARHLEGLLQKDRTKRLGAKDDFMEIKSHVFFSLINW 360

Db 301 NTAEMYDNILNKPLQKPNITNSARHLEGLLQKDRTKRLGAKDDFMEIKSHVFFSLINW 360

Qy 361 DDLINKKITPPFPNVSPGPNELRHFDPEFTEEPVNSIGKSPDSVLTASYKAAEAFLG 420

Db 361 DDLINKKITPPFPNVSPGPNELRHFDPEFTEEPVNSIGKSPDSVLTASYKAAEAFLG 420

Qy 421 FSYAPPTDSFL 431

Db 421 FSYAPPTDSFL 431

RESULT 6

AAW54025

ID AAW54025 standard; Protein: 431 AA.

XX AC AAW54025;

XX DT 31-JUL-1998 (first entry)

XX DE Human protein kinase HPK-3.

XX KW Protein kinase; human; HPK; signalling cascade; kinase expression;

XX KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;

XX KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;

XX OS Homo sapiens.

XX PN WO9811234-A2.

XX PD 19-MAR-1998.

XX PF 10-SEP-1997; 97WO-US15923.

PR 12-SEP-1996; 9605-0712709.
 XX (INCY-) INCYTE PHARM INC.
 PA Au-Young J, Guegler KJ, Hawkins PR;
 XX WPI; 1998-207394/18.
 DR N-PSDB; AAV23833.
 XX
 PT New isolated human protein kinase(s) - used to develop products for
 PT diagnosis and treatment of e.g. Alzheimer's disease,
 PT asthma, multiple sclerosis or rheumatoid arthritis
 XX
 PS Claim 23; Fig 3; 75pp; English.
 XX
 CC This sequence is a human protein kinase (HPK) of the invention. The
 CC HPK protein can be used to develop products for studying signalling
 CC cascades in various cells and tissues, diagnosing disease and selecting
 CC inhibitors or drugs with the potential to intervene in various disorders
 CC or diseases in which altered kinase expression is implicated. The
 CC products can be used to e.g. reverse memory loss such as due to
 CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
 CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
 CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
 CC lymphocytic leukaemias or lymphomas.
 XX
 SQ Sequence 431 AA;

Query Match 99.3%; Score 2254; DB 19; Length 431;
 Best Local Similarity 99.3%; Pred. No. 8.9e-203;
 Matches 428; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTVKTEAAKGLTYSRMGRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60
 DB 1 MAVKTEAAKGLTYSRMGRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60
 QY 61 SQQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKSGFGKVLARHKA 120
 DB 61 SQQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKSGFGKVLARHKA 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERNVLLKNVKHPFLVGLHFSFQADKLYFVLDYIN 180
 DB 121 EVFYAVKVLQKAILKKKEKHIMSERNVLLKNVKHPFLVGLHFSFQADKLYFVLDYIN 180
 QY 181 GGELEFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILLDSQGHIVLTD 240
 DB 181 GGELEFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILLDSQGHIVLTD 240
 QY 241 FGLCKENIEHNSTTTFQCTPEYLAPEVLHKKOPDYRTVDWVCLGAVLYEMLYGLPPFYSR 300
 DB 241 FGLCKENIEHNSTTTFQCTPEYLAPEVLHKKOPDYRTVDWVCLGAVLYEMLYGLPPFYSR 300
 QY 301 NTAEMYDNILNKPLQKPNITNSARHLLGGLQKDKRTKRLGAKDDFMEIKSHVFFSLINW 360
 DB 301 NTAEMYDNILNKPLQKPNITNSARHLLGGLQKDKRTKRLGAKDDFMEIKSHVFFSLINW 360
 QY 361 DDLINKKITPPPNVSGPNELRHDPDPEETEPVNSIGKSPDSVLVTASVKEAAEAFLG 420
 DB 361 DDLINKKITPPPNVSGPNELRHDPDPEETEPVNSIGKSPDSVLVTASVKEAAEAFLG 420
 QY 421 FSYAPPTDSFL 431
 DB 421 FSYAPPTDSFL 431

RESULT 7
 AAB24116
 ID AAB24116 standard; Protein; 431 AA.
 XX
 AC AAB24116;
 XX

DT 29-JAN-2001 (first entry)
 XX

DE Rat serum and glucocorticoid regulated kinase protein (RSGK).
 XX Human; STK3; serine threonine kinase; HSGK; RSGK;
 KW serum and glucocorticoid regulated kinase.
 XX Rattus sp.
 XX CN1259573-A.
 XX PD 12-JUL-2000.
 XX PF 29-OCT-1998; 98CN-0123822.
 XX PR 29-OCT-1998; 98CN-0123822.
 XX PA (UYFU-) UNIV FUDAN.
 XX PI Yu L, Fu Q, Zhao Y;
 XX WPI; 2000-587991/56.
 XX New human serine threonine protein kinase, its code sequence,
 PT preparation, and use -
 PS Example 3; Fig 4; 29pp; Chinese.
 XX
 CC The present invention describes human serine threonine kinase STK3,
 CC which is a new member of the serine/threonine protein kinase family.
 CC Human STK3 shares homology with human serum and glucocorticoid
 CC regulated kinase (HSGK) and rat SKK (RSGK). The present sequence
 CC represents the rat RSGK protein sequence from the present invention.
 XX
 SQ Sequence 431 AA;

Query Match 97.9%; Score 2222; DB 21; Length 431;
 Best Local Similarity 97.0%; Pred. No. 9e-200; 5; Indels 0; Gaps 0;
 Matches 418; Conservative 8; Mismatches 5;

QY 1 MTVKTEAAKGLTYSRMGRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60
 DB 1 MTVKTEAAKGLTYSRMGRGVAIIAFMKQRRMGLNDFIQKIANNYSACKHPEVQSILKI 60
 QY 61 SQQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKSGFGKVLARHKA 120
 DB 61 SQQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKSGFGKVLARHKA 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERNVLLKNVKHPFLVGLHFSFQADKLYFVLDYIN 180
 DB 121 EVFYAVKVLQKAILKKKEKHIMSERNVLLKNVKHPFLVGLHFSFQADKLYFVLDYIN 180
 QY 181 GGELEFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILLDSQGHIVLTD 240
 DB 181 GGELEFYHLQRCFCLEPRARFYAAEIASALGYLHSLNIVYRDLKPENILLDSQGHIVLTD 240
 QY 241 FGLCKENIEHNSTTTFQCTPEYLAPEVLHKKOPDYRTVDWVCLGAVLYEMLYGLPPFYSR 300
 DB 241 FGLCKENIEHNSTTTFQCTPEYLAPEVLHKKOPDYRTVDWVCLGAVLYEMLYGLPPFYSR 300
 QY 301 NTAEMYDNILNKPLQKPNITNSARHLLGGLQKDKRTKRLGAKDDFMEIKSHVFFSLINW 360
 DB 301 NTAEMYDNILNKPLQKPNITNSARHLLGGLQKDKRTKRLGAKDDFMEIKSHVFFSLINW 360
 QY 361 DDLINKKITPPPNVSGPNELRHDPDPEETEPVNSIGKSPDSVLVTASVKEAAEAFLG 420
 DB 361 DDLINKKITPPPNVSGPNELRHDPDPEETEPVNSIGKSPDSVLVTASVKEAAEAFLG 420
 QY 421 FSYAPPTDSFL 431
 DB 421 FSYAPPTDSFL 431

RESULT 8
 AAY93530

ID AAY93530 standard; Protein: 431 AA.
 XX
 AC RAY93530;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE A rat serum and glucocorticoid induced protein kinase.
 XX
 KW Protein kinase: Pkh1; Pkh2; Ypk1; Yrk2; protein kinase B-alpha;
 KW serum and glucocorticoid induced protein kinase; SGK; PKBalpha;
 KW 3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
 KW thrush; cancer; diabetes; obesity; antifungal; Candida infection.
 XX
 OS Rattus sp.
 XX
 PN W0200036135-A2.
 XX
 PD 22-JUN-2000.
 XX
 PF 14-DEC-1999; 99WO-GR04228.
 XX
 PR 14-DEC-1998; 98US-0112114.
 XX
 PA (MEDI-) MEDICAL RES COUNCIL.
 PA (REGC) UNIV CALIFORNIA.
 PI
 PI Thorner JW, Alessi DR, Torrance PD, Casamayor A;
 XX
 XX WPI: 2000-442391/38.
 DR N-PSDB; AAA46592.
 XX
 XX Screening method identifying compounds which modulate protein kinase
 PT activity for use in treating fungal infections and cancer -
 PT
 XX Disclosure; Fig 13; 155pp: English.
 XX
 PS The present sequence represents a rat serum and glucocorticoid induced
 CC protein kinase (SGK). The specification describes a screening method
 CC to identify a compound which modulates the activity of protein kinases
 CC from different sources, using host yeast cells. The method is used to
 CC identify a compound which modulates (inhibits) the activity of a
 CC protein kinase. Pkh1 or Pkh2 phosphorylate and activate Ypk1, Yrk2,
 CC SGK or protein kinase B-alpha (PKBalpha). 3-phosphoinositide-dependent
 CC protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
 CC Yrk2 or SGK but not PKBalpha or p70S6 kinase. Compounds identified by
 CC the methods are used to treat fungal infections e.g. thrush, and to
 CC treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the
 CC activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be
 CC used in the treatment of diabetes or obesity, and compounds which
 CC inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK
 CC (Ypk1 or Yrk2) can be used as an antifungal agent to treat Candida
 CC infections, e.g. thrush.
 XX
 SQ Sequence 431 AA:
 Query Match 97.9%; Score 2222; DB 21; Length 431;
 Best Local Similarity 97.0%; Pred. No. 9e-200;
 Matches 418; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MTVKTEAAKGTLYTSMRGWVALLIAFMKORRGLNDFIQTIANNSACKHPVEQSILKI 60
 DB 1 MTVKTEAAKSTLYTSMRGWVALLIAFMKORRGLNDFIQTIANNSACKHPVEQSYLKI 60
 QY 61 SQQPELNMNANPPSPSQQINLGPSSNPHAKPSDFHFLKVIKIGSGKVLARHKA 120
 DB 61 SQQPELNMNANPPSPSQQINLGPSSNPHAKPSDFHFLKVIKIGSGKVLARHKA 120
 QY 121 EYFYAVKVLQKKAILKKKEKHIMSERNVLLKNVHPFLVGLHFSFOTADKLYFLVDYIN 180
 DB 121 EYFYAVKVLQKKAILKKKEKHIMSERNVLLKNVHPFLVGLHFSFOTADKLYFLVDYIN 180
 QY 181 GGELEFVHLQRCFLPEPRARFYAAEIASALGYLHSLNIYVRDLKPENILLDSQGHIVLTD 240
 DB 181 GGELEFVHLQRCFLPEPRARFYAAEIASALGYLHSLNIYVRDLKPENILLDSQGHIVLTD 240
 QY 241 FGLCKENIEHNSTTSTFCUTPEYLAPEVLHKQPYDRTYDWCGLGAVLYEMLYGLPPFYSR 300
 DB 241 FGLCKENIEHNSTTSTFCUTPEYLAPEVLHKQPYDRTYDWCGLGAVLYEMLYGLPPFYSR 300
 QY 301 NTAEMVDNLNKLPLQKPNITNSARHLLLEGLOKDKTRKLGAKDDFMEIKSHVFTSLINW 360
 DB 301 NTAEMVDNLNKLPLQKPNITNSARHLLLEGLOKDKTRKLGAKDDFMEIKSHVFTSLINW 360
 QY 361 DOLINKKITPPFNPNVNSGPNELRHFDPEFTPEVPNSIGKSPDSVLVTASVKEAAAFIAG 420
 DB 361 DOLINKKITPPFNPNVNSGPNELRHFDPEFTPEVPNSIGKSPDSVLVTASVKEAAAFIAG 420
 QY 421 FSYAPPTDSFL 431
 DB 421 FSYAPPMDSFL 431
 RESULT 9
 AAB65614
 ID AAB65614 standard; Protein: 430 AA.
 XX
 AC AAB65614;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Novel protein kinase, SEQ ID NO: 140.
 XX
 KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.
 XX
 OS Mus musculus.
 XX
 PN W0200073469-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US14842.
 XX
 PR 28-MAY-1999; 99US-0136503.
 XX
 PA (SUGP-) SUGEN INC.
 PI Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX
 XX WPI: 2001-032161/04.
 DR N-PSDB; AAF44640.
 XX
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 PS Claim 10; Fig 1; 310pp: English.
 XX
 CC The present sequence is a novel protein kinase. The novel protein kinases
 CC and the nucleic acids that encode them may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure, bowel
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,

CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.

XX
 SQ Sequence 430 AA;
 Query Match 94.7%; Score 2150.5; DB 22; Length 430;
 Best Local Similarity 94.9%; Pred. No. 4.6e-193;
 Matches 410; Conservative 10; Mismatches 9; Indels 3; Gaps 2;
 QY 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRRMGNDLFIQIANNYSACKHPVQSILKI 60
 DB 1 MTVTEAAKGLTYSRMGMVAIIAFMKQRRMGNDLFIQIANNYSACKHPVQSILKI 60
 QY 61 SQPQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
 DB 61 SHPQPELMNANPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGKVLARHAE 120
 QY 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFYLDYIN 180
 DB 121 EVFYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADKLYFYLDYIN 180
 QY 181 GGELEYHLQRCFCLEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHVLTLD 240
 DB 181 GGELEYHLQRCFCLEPRARFYAAETASALGYLHSLNIVYRDLPENILLDSQGHVLTLD 240
 QY 241 -FGLCENIEHNTSTFCGTPPEYLAPEVLHKKQYDRTVDMWCLGAVLYEMLYGLPPFYS 299
 DB 241 XQL--RRIEHNNTSTFCGTPPEYLAPEVLHKKQYDRTVDMWCLGAVLYEMLYGLPPFYS 298
 QY 300 RNTAEYDNLKPLQKPNITNSARHLEGLLQKDRKRLGAKDDFMEIKSHVFFSLIN 359
 DB 299 RNTAEYDNLKPLQKPNITNSARHLEGLLQKDRKRLGAKDDFMEIKSHVFFSLIN 358
 QY 360 WDDLKNTTTPFNPNVSGPNELRHDFEETPEEPVNSIGKSPDSVLVTASVKEAAEAPL 419
 DB 359 WDDLKNTTTPFNPNVSGPNELRHDFEETPEEPVNSIGKSPDSVLVTASVKEAAEAPL 418
 QY 420 GFSYAPPTDSFL 431
 DB 419 GFSYAPPTDSFL 430

RESULT 10
 AAY48573
 ID AAY48573 standard; Protein; 388 AA.
 AC AAY48573;
 XX
 DT 08-DEC-1999 (first entry)
 DE Human breast tumour-associated protein 34.
 XX
 DE Expressed sequence tag; EST; human; breast; cancer; gene therapy;
 KW treatment; tumour; cytostatic; medicament.
 XX

OS Homo sapiens.
 XX
 PN DE19813839-A1.
 XX
 PD 23-SEP-1999.
 XX
 PF 20-MAR-1998; 98DE-1013839.
 XX
 PR 20-MAR-1998; 98DE-1013839.
 XX
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;
 XX
 DR WPI; 1999-528981/45.
 XX
 DR N-PSDB; AA233647.
 XX

PT Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -
 XX
 PS Claim 22; 157-158; 188pp; German.

XX
 CC This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AAY48540-Y48617 represent protein
 CC fragments encoded by the expressed sequence tags described in the method
 CC of the invention.

SQ Sequence 388 AA;

Query Match 87.6%; Score 1988; DB 20; Length 388;
 Best Local Similarity 98.9%; Pred. No. 7.3e-178;
 Matches 376; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 52 PEVQSIKISQPEELMANNPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGK 111
 DB 9 PEVQSIKISQPEELMANNPSPSPSQOINLGPSSNPHAKPSDFHFLKVIKGSFGK 68
 QY 112 VLLARHKAEEVYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADK 171
 DB 69 VLLARHKAEEVYAVKVLQKAILKKKEKHIMSERVLLKNVKKHPELVGLHFSFQADK 128
 QY 172 LFVLDYINGGELFYHLQRCFCLEPRARFYAAETASALGYLHSLNIVYRDLPENILLD 231
 DB 129 LFVLDYINGGELFYHLQRCFCLEPRARFYAAETASALGYLHSLNIVYRDLPENILLD 188
 QY 232 SGCHIVLTDGFLCKENIEHNTSTFCGTPPEYLAPEVLHKKQYDRTVDMWCLGAVLYEM 291
 DB 189 SGCHIVLTDGFLCKENIEHNTSTFCGTPPEYLAPEVLHKKQYDRTVDMWCLGAVLYEM 248
 QY 292 YGLPPFYSRNTAEYDNLKPLQKPNITNSARHLEGLLQKDRKRLGAKDDFMEIKS 351
 DB 249 YGLPPFYSRNTAEYDNLKPLQKPNITNSARHLEGLLQKDRKRLGAKDDFMEIKS 308
 QY 352 HVFFSLINWDDLKNTTTPFNPNVSGPNELRHDFEETPEEPVNSIGKSPDSVLVTASV 411
 DB 309 HVFFSLINWDDLKNTTTPFNPNVSGPNELRHDFEETPEEPVNSIGKSPDSVLVTASV 368
 QY 412 KEAAEAFSGSYAPPTDSFL 431
 DB 369 KEAAEAFSGSYAPPTDSFL 388

RESULT 11
 AAB99815
 ID AAB99815 standard; Protein; 373 AA.
 AC AAB99815;
 XX
 DT 20-SEP-2001 (first entry)
 DE AGC protein kinase family member SGK protein sequence.
 XX
 DE Protein kinase; identification; hydrophobic pocket; interacting;
 KW cancer; diabetes; inhibition; apoptosis; tissue injury;
 KW ischaemic injury; stroke.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN W020014497-A2.
 XX
 PD 21-JUN-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04598.
 XX

XX Modulation of serum and glucocorticoid-induced protein kinase (SGK) by
PT phosphorylation with 3-phosphoinositide-dependent protein kinase-1
PT (PK1) or dephosphorylation, useful for treatment of cancer, diabetes
PT and ischemic diseases -
XX
XX
XX Disclosure; Page 7; 127pp; English.
XX
XX The present sequence is that of human serum and
CC glucocorticoid-induced protein kinase (SGK) isoform 3. SGK
CC (see AAY9279) was initially identified as a glucocorticoid and
CC osmotic stress-responsive gene. SGK3 is activated by
CC phosphorylation in a similar manner to SGK. The invention provides
CC methods of activating SGK activity by phosphorylation using
CC 3-phosphoinositide-dependent protein kinase-1 (PK1), and of
CC reducing the activity of SGK by dephosphorylation. The invention
CC also provides a method of identifying a compound that modulates the
CC activity of SGK. Such compounds are useful for treating patients
CC requiring modulation of SGK, such as patients with cancer, diabetes
CC or ischaemic disease.
XX
XX Sequence 429 AA;
SQ
Query Match 64.88; Score 1472; DB 21; Length 429;
Best Local Similarity 67.28; Pred. No. 2.4e-129;
Matches 279; Conservative 53; Mismatches 69; Indels 14; Gaps 3;
QY 27 FMKORRMGLNDFIQKIANNSTYACKHPEVQSILKISQPELMNANSPPPP----- 77
DB 19 FIKORRAGNEFIQNLVRYPELYNPDVRAFLQMDSPKH---QSDPSEDEDESSQKHL 74
QY 78 SPQOQINLGPSNPHAKPSOPHFLKVIKGSFGKVLARHKAEEVFAVKVLQKAILKK 137
DB 75 STSQNIINLGPSGNPHAKPTDFDLKVIKGSFGKVLAKRKLKDGKFAVKVLQKIVLNR 134
QY 138 KEKHIMSERNVLLKNVKKHPELVGLHFSFOTADKLYFVLDYINGGELFYHLORERCLEP 197
DB 135 KEQKHIAERNVLLKNVKKHPELVGLHYSFOTELKLYFVLDVNGGELFFHLORERSFPEH 194
QY 198 RARYAABIASALGYLSLNTIVYRDLPENILSDSQHIVLTDGLCKENIEHNSTSTP 257
DB 195 RARYAABIASALGYLSIKIVYRDLPENILSDSVGHVLTDFGLCKEGIAISDTTTF 254
QY 258 CGTEYLAPVYLHKKOPVDTWMCGLGAVLYEMLYGLPPFYSRNTAEYDNLAKPLQLK 317
DB 255 CGTEYLAPVYRKOPYDNTVDMWCLGAVLYEMLYGLPPFYSRNTAEYDNLAKPLSLR 314
QY 318 PNITNSARHLLLEGLQKDRTRKLGAKDDFMKSHVFFSLINWDDLINKKITPPFPNVS 377
DB 315 PGVSLTAWSLLEELLEKDRQNLGAKEDFLIQNHPPFESLSWADLYQKKITPPFPNVA 374
QY 378 GPNELRHDPDETEPVPNSICKSPDSVLVTASVKEAAEFLGFSYAPPT-DSTL 431
DB 375 GPDDIRNPDFTAEETVPYSCVSDSYNASVLEADDAFVGFSPAPPSSEDLFL 429
RESULT 15
AAU87228
ID AAU87228 standard; Protein; 496 AA.
XX
XX AAU87228;
XX
XX 05-JUN-2002 (first entry)
XX
XX DE Novel central nervous system protein #138.
XX
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;

KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
XX food additive; food preservative; gene therapy.
OS Homo sapiens.
XX
XX WO200155318-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01332.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225277.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226868.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
XX 14-SEP-2000; 2000US-0233083.
XX 14-SEP-2000; 2000US-0233084.
XX 14-SEP-2000; 2000US-0233065.
XX 21-SEP-2000; 2000US-0234223.

[illegible]

pr me

sh.

much (con)hichartst di ss (A)nanasyn on, His scuc ic ed o b eve on, es wood illi

172

•

at c

WCS

WRA

JKG

5KG

১৩৬

03

NEC

230

TABLE

34

MJC

DEF:

21

15

Tue Aug 12 10:55:50 2003

us-10-000-039a-2.rag

Page 12

Search completed: August 11, 2003, 08:18:51
Job time : 139.884 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run On: August 11, 2003, 08:24:50 ; Search time 3332 Seconds
(without alignments)
1461.058 Million cell updates/sec

Title: US-10-000-039a-2_COPY_313_431

Perfect score: 624

Sequence: 1 PLQKPNITNSARHLLEGLL.....KEAAEAFGLFSYAPPTDSFL 119

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+p2n.model -DEV-xlp
-O/Ccp2_1/USPTQ.spool/US10000039/runat_01082003_085418_27511/app_query.fasta_1.661
-DB-GenEmbl -QFMT-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE-LOCAL
-OUTFMT-ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039.ccgcn_1_15569_runat_01082003_085418_27511 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hgt_hum.*
31: em_hgt_inv.*
32: em_hgt_other.*
33: em_hgt_mus.*
34: em_hgt_pln.*
35: em_hgt_rod.*
36: em_hgt_mam.*
37: em_hgt_vrt.*
38: em_sy.*
39: em_hgtgo_hum.*
40: em_hgtgo_mus.*
41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	624	100.0	1296	6	AX056374 Sequence
2	624	100.0	2370	6	AR179441 Sequence
3	624	100.0	2370	6	AX002570 Sequence
4	624	100.0	2370	6	AX337834 Sequence
5	624	100.0	2370	6	AX411211 Sequence
6	624	100.0	2370	9	Y10032 H sapiens m
7	621	99.5	2281	6	AX017284 Sequence
8	621	99.5	2281	6	AX524988 Sequence
9	621	99.5	2281	6	BD134455 Human nuc
10	621	99.5	2311	6	AR151390 Sequence
11	621	99.5	2311	6	AR270209 Sequence
12	621	99.5	2311	6	BD021928 Human pro
13	621	99.5	2354	6	AX553549 Sequence
14	621	99.5	2364	9	AK098509 Homo sapi
15	621	99.5	2382	9	AF153609 Homo sapi
16	621	99.5	2382	9	BC001263 Homo sapi
17	621	99.5	3196	6	AX713538 Sequence
18	621	99.5	3196	9	AK055077 Homo sapi
19	601	96.3	1659	10	BC002222 Mus muscu
20	601	96.3	2423	10	BC005720 Mus muscu
21	601	96.3	2426	10	AF205855 Mus muscu
22	601	96.3	2429	10	AF139638 Mus muscu
23	601	96.3	2432	6	AX056375 Sequence
24	582.5	93.3	1920	4	AF139639 Oryctolag
25	582.5	93.3	2435	10	RATSGPK
26	565	90.5	1417	5	AF057138 Xenopus l
27	517	82.9	2339	5	BC052134 Danio rer
28	517	82.9	2470	5	SAC223715 Squalus a
29	517	82.9	3105	5	SAC223716 Squalus a
30	492.5	78.9	113673	9	AL135839 Human DNA
31	479.5	76.8	5718	9	HSSGK
32	467.5	74.9	215058	2	AC103256 Homo sapi
33	459	73.6	142609	2	AC118095 Rattus no
34	459	73.6	240716	2	AC112280 Rattus no
35	452.5	72.5	208405	2	AC114405 Mus muscu
36	443.5	71.1	183140	10	AC124486 Mus muscu
37	367.5	58.9	1346	6	AX056376 Sequence
38	367.5	58.9	1812	6	AX056472 Sequence
39	367.5	58.9	1812	6	AX399684 Sequence
40	367.5	58.9	1812	9	AF169034 Homo sapi
41	367.5	58.9	1815	9	BC014037 Homo sapi
42	367.5	58.9	1834	6	AX026166 Sequence
43	367.5	58.9	2146	6	AX026164 Sequence
44	367.5	58.9	2146	9	AF186470 Homo sapi
45	362.5	58.1	2184	10	BC026549 Mus muscu

ALIGNMENTS

RESULT 1

AX056374 1296 bp DNA linear PAT 13-JAN-2001
LOCUS
DEFINITION Sequence 18 from Patent WO0073469.
ACCESSION AX056374
VERSION AX056374.1 GI:12229081
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Plowman,G.D., Martinez,R., Whyte,D. and Sudersanam,S.
TITLE Protein kinases
JOURNAL Patent: WO 0073469-A 18 07-DEC-2000;
Sugen, Inc. (US)
FEATURES
source Location/Qualifiers
1..1296
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 345 a 333 c 293 g 325 t
ORIGIN
Alignment Scores:
Pred. No.: 1.88e-65 Length: 1296
Score: 624.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-2_COPY_313_431 (1-119) x AX056374 (1-1296)
QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
Db 937 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCGAGGGCCTCCTG 996
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspAspPheMetGluIleLysSerHis 40
Db 997 CAGAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAAGATTAAAGATCAT 1056
QY 41 ValPhePheSerLeuIleAsnTrpAspSerValLeuValThrAlaSerValLys 60
Db 1057 GTCCTCTCTCTTAATTAACCTGGATGATCTCAATTAAGACAGATTAATCCGCCCTTT 1116
QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db 1117 AACCCAAATGTGAGTGGGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 1176
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 1177 CCGTGTCCCACTCCATTGGCAAGTCCCTCGACAGGCTCTCTGTACAGCCGCGTCAG 1236
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 1237 GAAGCTGCCGAGGCTTTCTAGGCTTTCTATGCGCTTCCTCCGACGAGCTTTCTCTC 1293
RESULT 2
AR179441
LOCUS
DEFINITION Sequence 1 from patent US 6326181.
ACCESSION AR179441
VERSION AR179441.1 GI:20220996
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Lang,F. and Waldegger,S.
TITLE Cell volume regulated human kinase h-sgk
JOURNAL Patent: US 6326181-A 1 04-DEC-2001;
FEATURES
source Location/Qualifiers
1..2370
Unclassified.
/organism="Homo sapiens"
/mol_type="genomic DNA"
/strain="CAUCASIAN"
/db_xref="taxon:9606"
/clone="HSRNASPK"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-65 Length: 2370
Score: 624.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-2_COPY_313_431 (1-119) x AX002570 (1-2370)

BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-65 Length: 2370
Score: 624.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-2_COPY_313_431 (1-119) x AR179441 (1-2370)
QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
Db 979 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCGAGGGCCTCCTG 1038
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspAspPheMetGluIleLysSerHis 40
Db 1039 CAGAAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAAGATTAAAGATCAT 1098
QY 41 ValPhePheSerLeuIleAsnTrpAspSerValLeuValThrAlaSerValLys 60
Db 1099 GTCCTCTCTCTTAATTAACCTGGATGATCTCAATTAAGACAGATTAATCCGCCCTTT 1158
QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db 1159 AACCCAAATGTGAGTGGGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 1218
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 1219 CCGTGTCCCACTCCATTGGCAAGTCCCTCGACAGGCTCTCTGTACAGCCGCGTCAG 1278
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 1279 GAAGCTGCCGAGGCTTTCTAGGCTTTCTATGCGCTTCCTCCGACGAGCTTTCTCTC 1335
RESULT 3
AX002570
LOCUS
DEFINITION Sequence 1 from Patent EP0861896.
ACCESSION AX002570
VERSION AX002570.1 GI:7242111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Lang,F. and Waldegger,S.D.
TITLE Cell volume regulated human kinase h-sgk
JOURNAL Patent: EP 0861896-A 1 02-SEP-1998;
DADE BEHRING MARBURG GMBH (DE)
FEATURES
source Location/Qualifiers
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/strain="CAUCASIAN"
/db_xref="taxon:9606"
/clone="HSRNASPK"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Alignment Scores:
Pred. No.: 3.9e-65 Length: 2370
Score: 624.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-000-039A-2_COPY_313_431 (1-119) x AX002570 (1-2370)

```

QY      1  ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
Db      979  CCTCTCCAGCTGAACCAATAATACAAATTCGCAAGACACCTCTCGAGGGCCCTCTG 1038
QY      21  GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db      1039  CAGAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGATCAT 1098
QY      41  ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
Db      1099  GTCTTCTCTCTTAATTAAGTGGATGATCTCTTAATAAGAGATTACTCCCTCTTTT 1158
QY      61  AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db      1159  AACCCAAATGTGAGTGGCCCAACGAGCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 1218
QY      81  ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db      1219  CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG 1278
QY      101  GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db      1279  GAAGCTCCGAGGCTTCTCTAGGCTTTCTCTATGCGCTCCACGAGACTCTTTCTCTC 1335

RESULT 4
AX337834
LOCUS      AX337834      2370 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 8343 from Patent WO0194629.
ACCESSION  AX337834
VERSION     AX337834.1  GI:18128553
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horridan,S., Soppet,D.R. and Weaver,Z.
TITLE       Cancer gene determination and therapeutic screening using signature
JOURNAL     Patent: WO 0194629-A 8343 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES    Location/Qualifiers
            source
            1..2370
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
BASE COUNT  636 a  517 c  513 g  704 t
ORIGIN
Alignment Scores:
Pred. No.:      3.9e-65      Length:      2370
Score:          624.00      Matches:    119
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:     0
DB:             6           Gaps:         0

US-10-000-039A-2_COPY_313_431 (1-119) x AX337834 (1-2370)
QY      1  ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
Db      979  CCTCTCCAGCTGAACCAATAATACAAATTCGCAAGACACCTCTCGAGGGCCCTCTG 1038
QY      21  GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db      1039  CAGAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGATCAT 1098
QY      41  ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
Db      1099  GTCTTCTCTCTTAATTAAGTGGATGATCTCTTAATAAGAGATTACTCCCTCTTTT 1158

US-10-000-039A-2_COPY_313_431 (1-119) x AX337834 (1-2370)
QY      61  AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db      1159  AACCCAAATGTGAGTGGCCCAACGAGCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 1218
QY      81  ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db      1219  CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG 1278
QY      101  GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db      1279  GAAGCTCCGAGGCTTCTCTAGGCTTTCTCTATGCGCTCCACGAGACTCTTTCTCTC 1335

RESULT 5
AX411211
LOCUS      AX411211      2370 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION Sequence 3858 from Patent WO0229103.
ACCESSION  AX411211
VERSION     AX411211.1  GI:21443916
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE       Gene expression profiles in liver cancer
JOURNAL     Patent: WO 0229103-A 3858 11-APR-2002;
GENE LOGIC INC (US)
FEATURES    Location/Qualifiers
            source
            1..2370
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /note="EMBL/GenBank Accession No. Y10032"
BASE COUNT  636 a  517 c  513 g  704 t
ORIGIN
Alignment Scores:
Pred. No.:      3.9e-65      Length:      2370
Score:          624.00      Matches:    119
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:     0
DB:             6           Gaps:         0

US-10-000-039A-2_COPY_313_431 (1-119) x AX411211 (1-2370)
QY      1  ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
Db      979  CCTCTCCAGCTGAACCAATAATACAAATTCGCAAGACACCTCTCGAGGGCCCTCTG 1038
QY      21  GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db      1039  CAGAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGATCAT 1098
QY      41  ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
Db      1099  GTCTTCTCTCTTAATTAAGTGGATGATCTCTTAATAAGAGATTACTCCCTCTTTT 1158
QY      61  AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db      1159  AACCCAAATGTGAGTGGCCCAACGAGCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 1218
QY      81  ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db      1219  CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG 1278
QY      101  GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db      1279  GAAGCTCCGAGGCTTCTCTAGGCTTTCTCTATGCGCTCCACGAGACTCTTTCTCTC 1335

RESULT 6
HSRNAS7PK

```

```

QY      61  AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db      1159  AACCCAAATGTGAGTGGCCCAACGAGCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 1218
QY      81  ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db      1219  CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG 1278
QY      101  GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db      1279  GAAGCTCCGAGGCTTCTCTAGGCTTTCTCTATGCGCTCCACGAGACTCTTTCTCTC 1335

RESULT 5
AX411211
LOCUS      AX411211      2370 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION Sequence 3858 from Patent WO0229103.
ACCESSION  AX411211
VERSION     AX411211.1  GI:21443916
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE       Gene expression profiles in liver cancer
JOURNAL     Patent: WO 0229103-A 3858 11-APR-2002;
GENE LOGIC INC (US)
FEATURES    Location/Qualifiers
            source
            1..2370
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /note="EMBL/GenBank Accession No. Y10032"
BASE COUNT  636 a  517 c  513 g  704 t
ORIGIN
Alignment Scores:
Pred. No.:      3.9e-65      Length:      2370
Score:          624.00      Matches:    119
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:     0
DB:             6           Gaps:         0

US-10-000-039A-2_COPY_313_431 (1-119) x AX411211 (1-2370)
QY      1  ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
Db      979  CCTCTCCAGCTGAACCAATAATACAAATTCGCAAGACACCTCTCGAGGGCCCTCTG 1038
QY      21  GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db      1039  CAGAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGATCAT 1098
QY      41  ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
Db      1099  GTCTTCTCTCTTAATTAAGTGGATGATCTCTTAATAAGAGATTACTCCCTCTTTT 1158
QY      61  AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db      1159  AACCCAAATGTGAGTGGCCCAACGAGCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 1218
QY      81  ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db      1219  CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG 1278
QY      101  GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db      1279  GAAGCTCCGAGGCTTCTCTAGGCTTTCTCTATGCGCTCCACGAGACTCTTTCTCTC 1335

RESULT 6
HSRNAS7PK

```


DEFINITION Sequence 39 from Patent EP1236799.

ACCESSION AX524988

VERSION AX524988.1 GI:25170070

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT 601 a 498 c 494 g 688 t

ORIGIN

Alignment Scores:

Pred. No.: 8,58e-65 Length: 2281

Score: 621.00 Matches: 118

Percent Similarity: 100.00% Conservative: 1

Best Local Similarity: 99.16% Mismatches: 0

Query Match: 99.52% Indels: 0

DB: 6 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x AX524988 (1-2281)

Qy 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20

Db 879 CCTCTCCAGCTGAACCAATATTACAAATTCGCGAGACACCTCTGGAGGCGCTCTG 938

Qy 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40

Db 939 CAGAAGGACAGGACAAAGCGCTCGGGGCCAAGATGACTTCATGGAGATTAAAGATCAT 998

Qy 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60

Db 999 GTCTTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1058

Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80

Db 1059 AACCCAAATGTAGTGGCGCCCAACGACCTAGCGACACTTTGACCCCGAGTTTACCGAAGAG 1118

Qy 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerVallys 100

Db 1119 CCTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGGCTCAAG 1178

Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119

Db 1179 GAAGTGGCGAGGCTTTCTAGGCTTTCTATGCGCTCCACGAGACTCTTTCCTC 1235

RESULT 9

BD134455

LOCUS

DEFINITION Human nucleic acid sequence originating in mammary tumor tissue.

ACCESSION BD134455

VERSION BD134455.1 GI:23229400

KEYWORDS JP 2002506643-A/33.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PR

PI

PI

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

```

source
1. .2311
/organism="unknown"
BASE COUNT 604 a 508 c 506 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8,71e-65 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x AR151390 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLeu 20
DB 946 CCTCTCCAGCTGAACCAATATTACAAATTCCTCCAGACACCTCTCTGGAGGCTCTG 1065

21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
DB 1006 CAGAAGGACAGCAAGCGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCA 1065

41 ValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysLysLysLysLys 60
DB 1066 GTCTTCTCTCTTAATTAATCTGGATGATCTCTATTAAAGAGATTACTCTCTCT 1125

61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
DB 1126 AACCCAAATGTAGTGGGCCCAAGGACCTACGGCCTTTCACCCGAGTTTACCGAAG 1185

81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerVal 100
DB 1186 CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCCGTCACAGCGAGCTCAAG 1245

101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraLapProThrAspSerPheLeu 119
DB 1246 GAAGCTGCGGAGGCTTCTTAGGCTTTTCTATGCGCTCCACGGACTCTTCTCTC 1302

RESULT 11
AR270209
LOCUS AR270209 2311 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 772 from patent US 6500938.
ACCESSION AR270209
VERSION AR270209.1 GI:29701443
KEYWORDS Location/Qualifiers
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Au-Young,J. and Seilhamer,J.J.
TITLE Composition for the detection of signaling pathway gene expression
JOURNAL Patent: US 6500938-A 772 31-DEC-2002;
FEATURES
source
1. .2311
/organism="unknown"
BASE COUNT 604 a 508 c 506 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8,71e-65 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x AR270209 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLeu 20
DB 946 CCTCTCCAGCTGAACCAATATTACAAATTCCTCCAGACACCTCTCTGGAGGCTCTG 1065

21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
DB 1006 CAGAAGGACAGCAAGCGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCA 1065

41 ValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysLysLysLysLys 60
DB 1066 GTCTTCTCTCTTAATTAATCTGGATGATCTCTATTAAAGAGATTACTCTCTCT 1125

61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
DB 1126 AACCCAAATGTAGTGGGCCCAAGGACCTACGGCCTTTCACCCGAGTTTACCGAAG 1185

81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerVal 100
DB 1186 CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCCGTCACAGCGAGCTCAAG 1245

101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraLapProThrAspSerPheLeu 119
DB 1246 GAAGCTGCGGAGGCTTCTTAGGCTTTTCTATGCGCTCCACGGACTCTTCTCTC 1302

RESULT 12
BD021928
LOCUS BD021928 2311 bp DNA linear PAT 27-AUG-2002
DEFINITION Human protein kinases.
ACCESSION BD021928
VERSION BD021928.1 GI:22563151
KEYWORDS JP 2001506843-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Young,J.O., Gegura,K.J. and Horkins,P.R.
TITLE Human protein kinases
JOURNAL Patent: JP 2001506843-A 3 29-MAY-2001;
INSITE PHARMACEUTICALS INC
COMMENT PN JP 2001506843-A/3
PD 29-MAY-2001
PF 10-SEP-1997 JP 1998513776
PR 12-SEP-1996 US 08/712709
PI JANICE O YOUNG, KARL J GEGURA, PHILLIP R HORKINS KC
C12N15/00,A61K38/45,A61K39/395,A61K39/395,A61K45/00,A61K48/00, PC
A61P11/06,
PC A61P19/02,A61P25/28,A61P29/00,A61P35/00,A61P43/00,C07K16/40,
C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/12,C12Q1/48,C12N15/00,C12N5/
PC 00,A61K37/52
CC
FEATURES
Key Location/Qualifiers.
FH
source
1. .2311
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 604 a 507 c 507 g 692 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 8,71e-65 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 6 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x BD021928 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLeu 20
DB 946 CCTCTCCAGCTGAACCAATATTACAAATTCCTCCAGACACCTCTCTGGAGGCTCTG 1005

21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
DB 1006 CAGAAGGACAGCAAGCGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCA 1065

```

QY	41	ValPheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProProPhe	60
Db	1066	GTCTTCTTCTCTTAATTAACGGGATGATCATTAATAAGAGATTACTCCCTCTTT	1125
QY	61	AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu	80
Db	1126	AACCCAAATGTGAGTGGGCCACAGACCTAGGACCTTTGACCCCGAGTTTACCGAAGAG	1185
QY	81	ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys	100
Db	1186	CCTGTCTCCCAACATTCGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG	1245
QY	101	GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu	119
Db	1246	GAAGCTGCCAGGCTTCTTAGGCTTTCTATGCGCTCCACGAGCTCTTCTCTC	1302
RESULT 13			
LOCUS	AX553549	2354 bp	DNA linear PAT 27-NOV-2002
DEFINITION	Sequence 1 from Patent WO02074987.		
ACCESSION	AX553549		
VERSION	AX553549.1	GI:25897549	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Lang, F., Busjahn, A. and Luft, F.C.		
TITLE	Quantitative diagnostic analysis of hypertension		
JOURNAL	Patent: WO 02074987-A 1 26-SEP-2002;		
Lang, Florian (DE)			
FEATURES	Location/Qualifiers		
source	1..2354		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	43..1338		
CDS	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD58123.1"		
	/db_xref="GI:25897550"		
	/translation="MTVKTEAAKGLTYSRMGRVAILIAPMKORRMGLNDFQKIAN		
	NSYAKPEVQSILKIQPELPMANPPSPSQINLSPNSPHAKPSDFHLK		
	VIGKSEFKVLLARHKAEEFYAVKVLQKAILKKEKHIMSERVLLKNVHPFLV		
	GLPFSFOTADKLVFVDYINGGELFYHLQRCERFLEPRARFYAAEIASALCYLHSLNI		
	YRDLPENILLDSQGHVLTDFGLCKENIEHNSTTTFCTGCTPEYLAPEVHLKQPYDR		
	TYDRLGCLNLYEMLYGLPPEYSRWAEMNDNLNKLPLQPKPNITNSARHLLEGLLQK		
	DRTRLGAKDDFMEIKSHVFFSLNWDLDLNLKLTTPFPNPNVGFNDLRFHDFPFTTE		
	PVPNSIGKSPDVLVTASVKEAAEAFILGFSYAPPTDSFL"		
variation	762		
	/note="1. SNP (C in T), stumme Mutation, d.h. beide		
	Versionen des SNPs resultieren in der AMINO ACID Asp in		
	der AMINO ACID- Position 240"		
BASE COUNT	613 a 518 c 513 g 710 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	8,91e-65	Length:	2354
Score:	621.00	Matches:	118
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.16%	Mismatches:	0
Query Match:	99.52%	Indels:	0
DB:	6	Gaps:	0
US-10-000-039A-2_COPY_313_431 (1-119) x AX553549 (1-2354)			
QY	1	ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu	20
Db	979	CCTCCACGCTGAACCAAAATATTACAAATTCGCGAAGACACCTCTCTGGAGCGCTCTG	1038
QY	21	GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis	40
Db	1039	CAGAAGCAGACAAAGCGGCTCGGGCCAAAGGATGACTTCTATGAGATTAAAGACTCAT	1098
QY	41	ValPheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProProPhe	60
Db	1099	GTCTTCTTCTCTTAATTAACGGGATGATCATTAATAAGAGATTACTCCCTCTTT	1158
QY	61	AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu	80
Db	1159	AACCCAAATGTGAGTGGGCCACAGACCTAGGACCTTTGACCCCGAGTTTACCGAAGAG	1218
QY	81	ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys	100
Db	1219	CCTGTCTCCCAACATTCGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAG	1278
QY	101	GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu	119
Db	1279	GAAGCTGCCAGGCTTCTTAGGCTTTCTATGCGCTCCACGAGCTCTTCTCTC	1335
RESULT 14			
LOCUS	AK098509	2364 bp	mRNA linear PRI 15-JUL-2002
DEFINITION	Homo sapiens cDNA FLJ25643 fis, clone STM07148, highly similar to		
ACCESSION	AK098509		
VERSION	AK098509.1	GI:21758535	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,		
	Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S.,		
	Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,		
	Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S.,		
	Morinaga, N., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,		
	Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isoqai, T. and		
	Sugano, S.		
	NEDO human cDNA sequencing project		
	Unpublished		
REFERENCE	2 (bases 1 to 2364)		
AUTHORS	Sugano, S. and Suzuki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,		
	University of Tokyo, Laboratory of Genome Structure, Human Genome		
	Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan		
	(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,		
	Fax:81-3-5449-5416)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of		
	Economy, Trade and Industry of Japan; cDNA full insert sequencing:		
	Research Association for Biotechnology (RAB); cDNA library		
	construction and 5'-end one pass sequencing: Institute of Medical		
	Science, University of Tokyo, Laboratory of Genome Structure, Human		
	Genome Center; 3'-end one pass sequencing: RAB; clone selection for		
	full insert sequencing: RAB and Helix Research Institute.		
FEATURES	Location/Qualifiers		
source	1..2364		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="STM07148"		
	/tissue_type="stomach mucosa"		
	/clone_lib="STM"		
	/note="cloning vector: pME18SFL3"		
BASE COUNT	615 a 524 c 517 g 708 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	8,96e-65	Length:	2364
Score:	621.00	Matches:	118
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.16%	Mismatches:	0
Query Match:	99.16%		

Query Match:	99.52%	Indels:	0
DB:	9	Gaps:	0
US-10-000-039a-2_COPY_313_431 (1-119) x AK098509 (1-2364)			
Qy	1	ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu	20
Db	994	CTCTCCAGCTGAACCAATATTACAATTCGCGAAGACACCTCCTCGAGGGCCCTCTG	1053
Qy	21	GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis	40
Db	1054	CAGAAGCACAGGACAAAGCGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCAT	1113
Qy	41	ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProProPhe	60
Db	1114	GTCTTCTCTCCCTTAATTAAGTGGATGATCTCATTAATTAAGAGATTACTCCCTCTTT	1173
Qy	61	AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu	80
Db	1174	AACCCAAATGTGAGTGGGCCCAACGACCTACGGCAGCTTTGACCCCGAGTTTACCGAAG	1233
Qy	81	ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys	100
Db	1234	CCTGTCCCAACTCCATGTCGAAGTCCCTGACAGGCTCTCGTCACAGCCAGCGTCAAG	1293
Qy	101	GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu	119
Db	1294	GAAGCTCCCGAGGCTTCTTCTAGGCTTTTCTATGGCGCTCCACGGAGCTCTTCTCTC	1350
RESULT 15			
AF153609	2382 bp	mRNA	linear
LOCUS	AF153609	Homo sapiens serine/threonine protein kinase sgk mRNA, complete cds.	PRI 28-JUN-1999
DEFINITION	AF153609	1 GI:5231142	
ACCESSION	AF153609	1	
VERSION	AF153609.1	GI:5231142	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
SOURCE			
CDS			

Search completed: August 11, 2003, 10:05:01
Job time : 3340 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run On: August 11, 2003, 08:23:34 ; Search time 443.545 Seconds
(without alignments)
724.239 Million cell updates/sec

Title: US-10-000-039A-2_COPY_313_431

Perfect score: 624

Sequence: 1 PLQKPNITNSARHLEGLL.....KEAAEAFGLFSYAPPTDSFL 119

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlip
-O=/cgn2_1/USPRO_spool/US10000039/runat_01082003_085418_27503/app_query.fasta_1.661
-DB=N.Geneseq_19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039 -CGN_1_1_740 -runat_01082003_085418_27503 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	624	100.0	1296	22	AAFA4639 Novel protein kina
2	624	100.0	2346	20	AAV74190 Human sqk DNA. Ho
3	624	100.0	2370	19	AAV48311 Human cell-volume
4	624	100.0	2370	24	ABK84749 Human cDNA differe
5	624	100.0	2370	24	ABN97360 Gene #3858 used to
6	624	100.0	2370	24	ABL70006 Pancreas cancer re
7	621	99.5	2281	20	AAZ33647 Human breast tumor
8	621	99.5	2311	19	AAV23833 Human protein kina
9	621	99.5	2311	25	ACA56174 Human signalling p
10	621	99.5	2343	25	ABF74395 Human cDNA sequenc
11	601	96.3	2432	22	AAFA4640 Novel protein kina
12	582.5	93.3	2365	21	AAA46592 DNA encoding a rat
13	512	82.1	404	25	ABX42261 Bovine EST associa
14	499	80.0	433	25	ABX38963 Bovine EST associa
15	367.5	58.9	1333	22	AAH99535 Human protein enco
16	367.5	58.9	1346	22	AAFA4641 Novel protein kina
17	367.5	58.9	1366	22	AAFA4987 cDNA encoding nove
18	367.5	58.9	1812	22	AAFA4737 Novel protein kina
19	367.5	58.9	1812	24	AAAD36141 Human serum and gl
20	367.5	58.9	1834	21	AAAZ37856 Human serum and gl
21	367.5	58.9	2146	21	AAAZ7857 Human serum and gl
22	360.5	57.8	2250	22	AAFA4642 Novel protein kina
23	360.5	57.8	2404	21	AAAZ7858 Human serum and gl
24	360.5	57.8	2483	22	AAK94684 Human serum glucoc
25	360.5	57.8	2512	20	AAV99653 Human serum glucoc
26	360.5	57.8	2572	23	ABK43712 DNA encoding novel
27	360.5	57.8	2702	23	ABK43558 Human polynucleoti
28	360.5	57.8	2711	22	AAI59776 Human polynucleoti
29	360.5	57.8	2760	22	AAI57990 Human serine threo
30	339.5	57.5	3019	21	AAA99248 cDNA encoding nove
31	330.5	56.2	851	22	AAAS27021 DNA encoding novel
32	330.5	56.2	851	23	ABK43988 Human colon specif
33	260.5	41.7	3872	24	ABA91312 Mouse ischaemic co
34	244	39.1	431	24	ABI99409 Drosophila melanog
35	233.5	35.8	2375	23	ABL19107 Drosophila melanog
36	223.5	35.8	2425	23	ABL28911 Androgen receptor
37	223.5	35.8	3712	25	ABQ76675 Human pancreatic c
38	208.5	33.4	615	24	ABV96134 Clone 12702 cDNA f
39	203.5	32.6	257	19	AAV33520 Lung cancer relate
40	201	32.2	2346	24	ABL65827 Lung p70 S6 kinas
41	201	32.2	2346	25	ABZ74945 Human breast tumor
42	193	30.9	316	22	AAH55574 Human Akt-3 coding
43	191.5	30.7	1440	21	AAAG2451 Human Akt-3 encodin
44	191.5	30.7	1440	24	ABL50839 Human Akt-3 unclon
45	191.5	30.7	1547	21	AAAG2450 Human Akt-3 unclon

ALIGNMENTS

RESULT 1

AAFA4639

ID AAFA4639 standard: cDNA: 1296 BP.

XX

AC AAFA4639;

XX

DT 27-MAR-2001 (first entry)

DE Novel protein kinase cDNA, SEQ ID NO: 18.

XX

Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic; immunosuppressive; cardiac; renal; antiinflammatory; antistimatic; dermatological; antidiabetic; antifertility; gene therapy; vaccine; immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

OS Homo sapiens.
 XX WO200073469-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000: 2000HO-US14842.
 XX 28-MAY-1999: 99US-0136503.
 XX (SUGF-) SUGEN INC.
 XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX WPI: 2001-032161/04.
 XX P-PSDB: AAB65613.
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 XX treating immune-related diseases and disorders, cardiovascular disease,
 XX neurodegenerative diseases and/or cancers.
 XX Disclosure; Fig 2: 310pp; English.
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 XX and the protein kinases they encode may be used in the treatment and
 XX diagnosis of diseases associated with inappropriate kinase expression
 XX such as immune-related diseases and disorders, cardiovascular disease,
 XX neurodegenerative diseases and/or cancers. The nucleic acids and
 XX complementary sequences may also be used as DNA probes in diagnostic
 XX assays. The kinase polypeptides may be used as antigens in the production
 XX of antibodies of kinase expression and activity. Anti-kinase antibodies
 XX and kinase antagonists may also be used to down regulate kinase
 XX expression and activity. Diseases related to kinase expression and
 XX activity include rheumatoid arthritis, atherosclerosis, autoimmune
 XX disorders, complications of organ transplantation, myocardial infarction,
 XX immune disorders, cardiomyopathies, strokes, renal failure,
 XX oxidative-stress related disorders, chronic inflammatory bowel disease,
 XX chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 XX osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 XX reproductive disorders.
 XX Sequence 1296 BP; 345 A; 333 C; 293 G; 325 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1.12e-73 Length: 1296
 Score: 624.00 Matches: 119
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-10-000-039A-2_COPY_313_431 (1-119) x AAF44639 (1-1296)
 QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLeu 20
 DB 937 CCTCTCCAGCTGAACCAATATATACAAATTCGCGAGACACCTCTCGAGGCGCTCTG 996
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
 DB 997 CAGAGGACAGGACAGACCGCTCGGGCCCGCAGATGATCTCATGAGATGATGATCAT 1056
 QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
 DB 1057 GTCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAAGTCTCTCTTT 1116
 QY 61 AsnProAsnValSerGlyProAsnGlnLeuArgHisPheAspProGluPheThrGlu 80
 DB 1117 AACCCAAATGTGATGGGCCCAACAGACGTACGGACACTTTCACCCGAGTTTACCGAGAG 1176
 QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 DB 1177 CCTGTCCCACTCCATGGCAAGTCCCTGACAGCGTCTCTCTCACAGCCACGCTCAAG 1236

QY 101 GluAlaAlaClnAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 1237 GAAGCTGGCGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCGACGACTCTTCTTC 1293

RESULT 2
 AAV74190
 ID AAV74190 standard; DNA: 2346 BP.
 XX AAV74190:
 XX 15-MAR-1999 (first entry)
 XX Human sgk DNA.
 XX Serum glucocorticoid regulated kinase; sgk; human; treatment; inhibitor;
 XX serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 XX chronic renal failure; inflammation; Alzheimers disease; wound; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 45..1340
 XX FT /*tag- a
 XX FT /product- "sgk"
 XX FT /transl_except- (pos:1185..1187, aa:Asp)
 XX EP887081-A2.
 XX 30-DEC-1998.
 XX 27-MAY-1998; 98EP-0304189.
 XX 27-JUN-1997; 97US-0051124.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX Kumar JM;
 XX WPI: 1999-047627/05.
 XX P-PSDB: AAW90139.
 XX Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 XX disease - by administration of nucleic acids and antagonists which
 XX inhibit activity or expression of human serum glucocorticoid
 XX regulated kinase (sgk), a serine/threonine protein kinase
 XX Disclosure; Page 14-15; 17pp; English.
 XX This sequence encodes a novel human serum glucocorticoid regulated kinase
 XX (sgk) protein which is a member of the serine/threonine protein kinase
 XX family. This protein is used for the treatment of a subject having need
 XX to inhibit/antagonise activity or expression of human sgk polypeptide
 XX e.g. for the treatment of chronic renal failure, diabetic nephropathy,
 XX inflammation, Alzheimers disease and wounds.
 XX Sequence 2346 BP; 608 A; 520 C; 518 G; 700 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 2.62e-73 Length: 2346
 Score: 624.00 Matches: 119
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0
 US-10-000-039A-2_COPY_313_431 (1-119) x AAV74190 (1-2346)
 QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLeu 20
 DB 981 CCTCTCCAGCTGAACCAATATATACAAATTCGCGAGACACCTCTCTGAGGCGCTCTG 1040
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40

Db 1041 CAGAGGACAGGACAAAGCGGCTCGGGCCAAAGGATGACTTCATGAGAGATTAAAGAGTCAT 1100
 Qy 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProProPhe 60
 Db 1101 GCTCTCTCTCTCTTAATTAACCTGGATGATCTCATTAATAAGAGATTACCTCCCTTT 1160
 Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 Db 1161 AACCCAAATGCTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 1220
 Qy 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 Db 1221 CTTGTCCTCCCACTCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 1280
 Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 Db 1281 GAAGCTGCCGAGGCTTTCTAGGGTTTCTCTATGCGCTCCACCGGACTCTTTCTCTC 1337

RESULT 3

AAV48311
 ID AAV48311 standard; cDNA; 2370 BP.

XX AAV48311;

XX AC

XX AC

XX 16-NOV-1998 (first entry)

XX Human cell-volume regulating kinase h-sgk.

XX ss; human; cell-volume; kinase; h-sgk; diabetes mellitus;

XX renal insufficiency; inflammation; Alzheimer's disease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 43..1338

XX FT /*cag= a

XX FT /product= "Kinase h-sgk"

XX PN EP861896-A2.

XX PD 02-SEP-1998.

XX 27-JAN-1998; 98EP-0101338.

XX 28-FEB-1997; 97DE-1008173.

XX (DADE-) DADE BEHRING MARBURG GMBH.

XX Lang F, Waldegger S;

XX WPI; 1998-449109/39.

XX P-ESDB; AAW77219.

XX New nucleic acid encoding cell-volume regulating kinase h-sgk and
 PT related proteins - used for diagnosis and treatment of diseases
 PT involving changes in cell volume, e.g. renal insufficiency,
 XX inflammation, infections etc.
 XX disclosure; Fig 1; 15pp; German.
 XX The human cell-volume regulating kinase h-sgk is inhibited by the
 CC swelling of cells (or presence of urea), whereas cell shrinkage
 CC stimulates its expression. The nucleic acid h-sgk, and fragments, are
 CC particularly used to detect changes in cell volume, specifically for
 CC diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 CC natriemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 CC hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 CC intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 CC The nucleic acid, protein and products including receptors that bind
 CC h-sgk, can be used to treat these disorders.
 XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:

Pred. No.: 2.66e-73 Length: 2370
 Score: 624.00 Matches: 119
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 19 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x AAV48311 (1-2370)

Qy 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 Db 979 CCTCTCCAGCTGAAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGCGCTCTCTG 1038
 Qy 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspAspPheMetGluLleLysSerHis 40
 Db 1039 CAGAAAGCAGCAGCAAGCGCTCGGGCGCAAGGATGACTTCATGAGAGATTAAAGAGTCAT 1098
 Qy 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProProPhe 60
 Db 1099 GTCTTCTCTCTTAATTAACCTGGATGATCTCATTAATAAGAGATTACCTCCCTTT 1158
 Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 Db 1159 AACCCAAATGCTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 1218
 Qy 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 Db 1219 CTTGTCCTCCCACTCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 1278
 Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 Db 1279 GAAGCTGCCGAGGCTTTCTAGGGTTTCTCTATGCGCTCCACCGGACTCTTTCTCTC 1335

RESULT 4

ABK84749
 ID ABK84749 standard; cDNA; 2370 BP.

XX AC

XX ABK84749;

XX 14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #1320.
 XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; AKOS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

XX OS

XX WO200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US30821.

XX 03-OCT-2000; 2000US-237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity -

XX PS Claim 1; SEQ ID NO 1320; 114pp; English.

XX CC The invention relates to detecting (M1) granulocyte (GC) activation

CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by

CC DNA chip analysis as given in the specification, and comparing

CC the expression level to an expression level in an unactivated

CC GC, where differential expression of Gs is indicative of GCA.

CC Also included are modulating (M2) Gs by contacting GC with an agent

CC that alters the expression of at least one gene in Gs; (2) screening (M3)

CC for an agent capable of modulating GCA or an inflammation (especially

CC chronic) in a tissue, an allergic response in a subject, exposure of a

CC subject to a pathogen or sterile inflammatory disease using the

CC gene expression profile; (3) detecting (M4) an inflammation (especially

CC chronic) in a tissue, an allergic response in a subject, exposure of a

CC subject to a pathogen or sterile inflammatory disease, by detecting the

CC level of expression in a sample of the tissue of gene(s) from Gs, where

CC the level of expression of the gene is indicative of inflammation;

CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,

CC an allergic response in a subject, exposure of a subject to a pathogen

CC or sterile inflammatory disease, by contacting a tissue having

CC inflammation with an agent that modulates the expression of gene(s)

CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for

CC modulating GCA; M3 is useful for screening an agent capable of modulating

CC GCA preferably in an inflammation in a tissue; M4 is useful for

CC detecting an inflammation (especially chronic) in a tissue, an allergic

CC response in a subject, exposure of a subject to a pathogen or sterile

CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,

CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal

CC reperfusion injury, ARDS, adult respiratory distress syndrome,

CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,

CC periodontal disease; also bacterial infection, viral infection,

CC parasitic infection, protozoal infection, fungal infection and M5 is

CC useful for treating one of the above conditions. The present

CC sequence represents a gene differentially expressed in granulocytes.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2370 BP: 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:

Pred. No.:	2,66e-73	Length:	2370
Score:	624.00	Matches:	119
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-000-039A-2_COPY_313_431 (1-119) x ABK84749 (1-2370)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20

DB 979 CCTCTCCAGCTGAACCAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTG 1038

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40

DB 1039 CAGAGGACAGACGACAAAGCGCTCGGGCCCAAGGATGCTCATGGAGATTAAGAGTCAT 1098

QY 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60

DB 1099 GTCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTCTTT 1158

QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80

DB 1159 AACCAANTGTGACTGGGCCCAAGCAGCTAGCGGCACTTGCACCCGAGTTTACCAAGAG 1218

QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100

DB 1219 CCGTGTCCCACTCCATTTGGCAAGTCCCTGACAGCGCTCGTCACAGCCAGCGTCAAG 1278

QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 119

DB 1279 GAAGCTCCGAGGCTTTCTTAGGCTTTTCCTATGCGCTCCCGGAGCTCTTTCTTC 1335

RESULT 5

ABN97360

ID ABN97360 standard; DNA; 2370 BP.

XX AC ABN97360;

DF 13-AUG-2002 (first entry)

XX Gene #3858 used to diagnose liver cancer.

DE Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US30589.

XX 02-OCT-2000; 2000US-237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-De-Silva S, Vockley JG;

XX WPI: 2002-426119/45.

XX Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

PT liver tissue sample

XX Claim 1; SEQ ID NO 3858; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver

CC tumour in a patient, and differentiating metastatic liver cancer from

CC hepatocellular carcinoma in a patient, involving detecting the level of

CC expression of two or more genes represented in ABN93503-ABN97455 in a

CC tissue sample. The method of the invention has hepatotropic, and

CC cytostatic activity. The method is useful for diagnosing and detecting

CC the progression of liver cancer, hepatocellular carcinoma and metastatic

CC liver carcinoma in a patient. The method is useful for identifying

CC expression profiles which serve as useful diagnostic markers as well as

CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 2370 BP: 636 A; 517 C; 513 G; 704 T; 0 other;

Alignment Scores:

Pred. No.:	2,66e-73	Length:	2370
Score:	624.00	Matches:	119
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-000-039A-2_COPY_313_431 (1-119) x ABN97360 (1-2370)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20

DB 979 CCTCTCCAGCTGAACCAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTG 1038

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40

Db 1219 CCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGGTCACAG 1278
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProthraspSerPheLeu 119
Db 1279 GAAGCTGCGAGGCTTTCTAGGCTTTTCTATGCGCTCCACAGGACTCTTTCCTC 1335
RESULT 7
AA233647
ID AA233647 standard; cDNA: 2281 BP.
XX AA233647;
AC AA233647;
DT 08-DEC-1999 (first entry)
XX
XX Human breast tumour-associated EST 37.
DE
DE Expressed sequence tag; EST: human; breast; cancer; gene therapy;
KW treatment; tumour; cytostatic; medicament; ss.
XX
XX Homo sapiens.
OS
XX DE19813839-A1.
PN
XX
XX 23-SEP-1999.
XX
XX 20-MAR-1998; 98DE-1013839.
PF
XX
XX 20-MAR-1998; 98DE-1013839.
PR
XX
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
PA
XX
XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosentahl A;
PI
XX WPI: 1999-528981/45.
DR
XX P-PSDB; AA48573.
DR
XX
XX Human nucleic acid sequences and protein products from tumor breast
PT
PT tissue, useful for breast cancer therapy -
PT
XX
XX Claim 1a: 116-117; 188pp; German.
XX
XX This invention describes novel human nucleic acid sequences from tumor
CC breast tissue which have cytostatic activity. The nucleic acid sequences
CC can be used to produce and isolate full-length gene sequences. They can
CC be used to express proteins, which can be used as tools to find an
CC activity against breast cancer. The sequences can be used in sense or
CC antisense form. They are especially useful for medicaments for gene
CC therapy to treat breast cancer. AA233647 represents expressed
CC sequence tags described in the method of the invention.
XX
XX Sequence 2281 BP; 601 A; 498 C; 494 G; 688 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 6,388-73 Length: 2281
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservativity: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 20 Gaps: 0
US-10-000-039a-2_copy_313_431 (1-119) x AA233647 (1-2281)
QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeu 20
Db 879 CCTCTCCAGCTGAACCAATATTACAAATTCGCGACACACCTCTCTGGAGGGCTCTCTG 938
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db 939 CAGAGGACGACGACGACGCGCTCCGGCCAGGAGTACTTCATGGAGATTAAGAGTAT 998
QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
Db 999 GTCTTCTCTCTTAATTAACCTGGGATGATCTCATTAATAAGAGATTACTCCCTCTTT 1058

QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
Db 1059 AACCCAAATGTAGTGGGCCACAGCTACCGGACATTTGACCCGAGTTTACCGAGAG 1118
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 1119 CCGTCCCAACTCCATTGGCAAGTCCGCTGACAGCGTCTCTGTCACAGCCAGGTCACAG 1178
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProthraspSerPheLeu 119
Db 1179 GAAGCTGCGAGGCTTTCTAGGCTTTTCTATGCGCTCCACAGGACTCTTTCCTC 1235
RESULT 8
AAV23833
ID AAV23833 standard; cDNA: 2311 BP.
XX AAV23833;
AC AAV23833;
XX
XX 31-JUL-1998 (first entry)
DT
XX Human protein kinase HPK-3 coding sequence.
DE
DE
XX Protein kinase; human; HPK; signalling cascade; kinase expression;
KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;
KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;
KW lymphoma; therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 379..1080
FT CDS /*tag= a
ET
XX
XX WO9811234-A2.
PN
XX
XX 19-MAR-1998.
PD
XX
XX 10-SEP-1997; 97WO-US15923.
PF
XX
XX 12-SEP-1996; 96US-0712709.
PR
XX
XX (INCY-) INCYTE PHARM INC.
PA
XX
XX Au-Young J, Guegler KJ, Hawkins PR;
PI
XX WPI: 1998-207394/18.
DR
XX P-PSDB; AA54205.
DR
XX
XX New isolated human protein kinase(s) - used to develop products for
PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
PT asthma, multiple sclerosis or rheumatoid arthritis
XX
XX Claim 25; Fig 3; 75pp; English.
XX
XX This sequence encodes a human protein kinase (HPK) of the invention. The
CC HPK protein can be used to develop products for studying signalling
CC cascades in various cells and tissues, diagnosing disease and selecting
CC inhibitors or drugs with the potential to intervene in various disorders
CC or diseases in which altered kinase expression is implicated. The
CC products can be used to e.g. reverse memory loss such as due to
CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
CC lymphocytic leukaemias or lymphomas.
XX
XX Sequence 2311 BP; 604 A; 508 G; 692 T; 1 other;
SQ
Alignment Scores:
Pred. No.: 6,5e-73 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservativity: 1
Best Local Similarity: 99.16% Mismatches: 0

Query Match: 99.52% Indels: 0
DB: 19 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x AAV23833 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
|||||
DB 946 CCTCTCCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCCG 1005

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLysSerHis 40
|||||
DB 1006 CAGNAGGACAGGACAAAGCGCTCGGCGCAAGATGACTTCATGGAGATTAGATCAT 1065

QY 41 ValPheSerLeuLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
|||||
DB 1066 GTCTTCCTCTCTTAATTAAGTGGATGATCTCATTAAAGAGATTACTCCCTCTTT 1125

QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
|||||
DB 1126 AACCAATGTGAGTGGCGCCACAGGACCTACGGCACTTTGACCGCGATTACCGAGAG 1185

QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
|||||
DB 1186 CCTGTCCCGACTCATGGCAAGTCCCTGACAGCGTCTCTGACAGCGCTCAAG 1245

QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
|||||
DB 1246 GAAGTGGCGAGGCTTCTAGGCTTTCTATGCGCTCCACAGGACTTTTCCTC 1302

RESULT 9
ACA56174
ID ACA56174 standard; cDNA: 2311 BP.
AC ACA56174;
XX
XX
DT 06-JUN-2003 (first entry)
DE Human signalling pathway polynucleotide probe SEQ ID NO 772.
XX
XX
KW Human; probe; ss; array element; parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
XX US500938-B1.
XX
XX 31-DEC-2002.
XX
XX 30-JAN-1998; 98US-0016434.
XX
XX 30-JAN-1998; 98US-0016434.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Au-Young J, Seilhamer JJ;
XX
XX WPI: 2003-352189/33.
XX
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides.
XX
XX Claim 1; SEQ ID NO 772; 65pp; English.
XX
XX
CC The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,

CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=0650938B1.
XX
SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;

Alignment Scores:
Pred. No.: 6,5e-73 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 25 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x ACA56174 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
|||||
DB 946 CCTCTCCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCCG 1005

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLysSerHis 40
|||||
DB 1006 CAGNAGGACAGGACAAAGCGCTCGGCGCAAGATGACTTCATGGAGATTAGATCAT 1065

QY 41 ValPheSerLeuLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
|||||
DB 1066 GTCTTCCTCTCTTAATTAAGTGGATGATCTCATTAAAGAGATTACTCCCTCTTT 1125

QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
|||||
DB 1126 AACCAATGTGAGTGGCGCCACAGGACCTACGGCACTTTGACCGCGATTACCGAGAG 1185

QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
|||||
DB 1186 CCTGTCCCGACTCATGGCAAGTCCCTGACAGCGTCTCTGACAGCGCTCAAG 1245

QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
|||||
DB 1246 GAAGTGGCGAGGCTTCTAGGCTTTCTATGCGCTCCACAGGACTTTTCCTC 1302

RESULT 10
ABX74395
ID ABX74395 standard; cDNA: 2343 BP.
XX
XX AC ABX74395;
XX
XX DT 21-MAR-2003 (first entry)
XX
XX Human cDNA sequence #6 up-regulated in non-aggressive CC-RCC.
XX
XX Human; microarray; solid surface; immobilised probe; CC-RCC;
XX differential expression profile; aggressive CC-RCC tumour type;
XX non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
XX gene expression profiling; tumour tissue; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200279411-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 29-MAR-2002; 2002WO-US09576.

XX PR 29-MAR-2001; 2001US-279411P.
 XX XX (VAND-) VAN ANDEL INST.
 XX PA Haab B, Rhodes D, Teh BT, Takashi M;
 XX PI WPI, 2003-040679/03.
 XX DR New microarray, comprising a matrix of cDNA probe from a set of probes
 XX PT immobilised to a solid surface in predetermined order, useful in the
 XX PT prognosis of patients with clear cell renal carcinoma.
 XX PS Claim 1; Page 79-80; 179pp; English.
 XX CC The present invention relates to a microarray comprising a matrix of
 XX CC at least one cDNA probe from a set of probes immobilised to a solid
 XX CC surface in a predetermined order, where a row of pixels corresponds
 XX CC to replicates of one distinct probe from the set. The probes are
 XX CC complementary to nucleic acid sequences that are expressed
 XX CC differentially in aggressive as compared to non-aggressive types of
 XX CC clear cell renal carcinoma (CC-RCC) and that hybridise to the probes
 XX CC under high stringency conditions. The microarray is useful for the
 XX CC prognosis of patients with CC-RCC, wherein aggressive and
 XX CC non-aggressive CC-RCC tumour types are characterised by differential
 XX CC expression profiles of genes that hybridise with one or more probes
 XX CC immobilised on the microarray. The arrays are useful for gene
 XX CC expression profiling of tumour and normal tissues. The present
 XX CC sequence represents a human cDNA sequence up-regulated in
 XX CC non-aggressive CC-RCC phenotypes.
 XX SQ Sequence 2343 BP; 612 A; 517 C; 511 G; 703 T; 0 other;
 Alignment Scores:
 Pred. No.: 6.63e-73 Length: 2343
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 25 Gaps: 0
 US-10-000-039a-2_COPY_313_431 (1-119) x ABX74395 (1-2343)
 QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 972 CCTCTCCAGCTGAACCAAAATATTACAAATTTCCCAAGACACCTCTCTGGAGGCCCTCCG 1031
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLysSerHis 40
 DB 1032 CAGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTACAGTCAT 1091
 QY 41 ValPhePheSerLeuIleAsnTrpAspLeuIleAsnLysLysIleThrProPhe 60
 DB 1092 GTCCTCTCTCTTAATTAACCTGGAGATCTCATTAAAGAGATTACTCCCTCTTT 1151
 QY 61 AsnProAsnValSerGlyProAsnGlnLeuArgHisPheAspProGluPheThrGlu 80
 DB 1152 RACCCAAATGTGAGTGGCCCAACAGCCTACGGCCTTTCACCCCGAGTTTACCGAAG 1211
 QY 81 ProValProAsnSerIleGlySerProAspSerValLeuValThrAlaSerValLys 100
 DB 1212 CCTGTCCCACTCATTTGGCAAGTCCCTGACAGCTCTCTGTCACAGCCGCAAG 1271
 QY 101 GluAlaAlaGluAlaPheSerTyraIaProProThrAspSerPheLeu 119
 DB 1272 GAAGCTGCGGAGCTTTCTTAGGCTTTTCTATCGGCTCCACGAGACTCTTTCTC 1328
 RESULT 11
 AAF44640
 ID RAF44640 standard; cDNA; 2432 BP.
 XX XX
 AC AAF44640;
 XX XX

DT 27-MAR-2001 (first entry)
 XX DE Novel protein kinase cDNA. SEQ ID NO: 19.
 XX KW Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic;
 KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KW immune disorder; cardiovascular disease; neurodegenerative disease;
 KW cancer; autoimmune disease; stroke; inflammatory bowel disease;
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX OS Mus musculus.
 XX PN WO200073469-A2.
 XX XX 07-DEC-2000.
 XX PD 26-MAY-2000; 2000WO-US14842.
 XX PF 28-MAY-1999; 99US-0136503.
 XX PR (SUG-) SUGEN INC.
 XX PA Plowman GD, Martinez R, Whyte D, Sudersanam S;
 PI WPI; 2001-022161/04.
 XX DR P-FSDB; RA65614.
 XX XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers.
 XX PS Disclosure: Fig 2; 310pp; English.
 XX CC The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes and
 CC reproductive disorders.
 XX SQ Sequence 2432 BP; 586 A; 586 C; 560 G; 699 T; 1 other;
 Alignment Scores:
 Pred. No.: 3.41e-70 Length: 2432
 Score: 601.00 Matches: 112
 Percent Similarity: 99.16% Conservative: 6
 Best Local Similarity: 94.12% Mismatches: 1
 Query Match: 96.31% Indels: 0
 DB: 22 Gaps: 0
 US-10-000-039a-2_COPY_313_431 (1-119) x AAF44640 (1-2432)
 QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
 DB 1008 CCTCTCCAGCTGAACCAAAATATTACAAATTTCCCAAGACACCTCTCTGGAGGCCCTCCG 1067
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLysSerHis 40
 DB 1068 CAGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTACAGTCAT 1127
 QY 41 ValPhePheSerLeuIleAsnTrpAspLeuIleAsnLysLysIleThrProPhe 60

DB 1128 ATTTCTCTCTTTATTAACCTGGGATGATCTCATCAATAGAGATTACACCCCATTT 1187
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
 DB 1188 AACCCAAATGTAGTGGGCCAGTGCACCTTCGACATTCGATCCGAGTTACCGAGGAG 1247
 QY 81 ProValProAsnSerTleGlyLysSerProAspSerValLeuValThrAlaSerVallys 100
 DB 1248 CCGTCCCGAGCTCCATCGCAGGTCCTGACAGCATCTTGTACCGCCAGTGTGAAG 1307
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 1308 GAGCAGCAGAGAGCCCTCTCGCTCTCTATGACACCTCCCTGTGTGATCTCTCTC 1364

RESULT 12

AAA46592
 ID AAA46592 standard; DNA; 2365 BP.

XX AC AAA46592;

XX DT 25-SEP-2000 (first entry)

XX DNA encoding a rat serum and glucocorticoid induced protein kinase.

XX Protein kinase; Pkh1; Pkh2; Ypk1; Yrk2; protein kinase B-alpha;
 KW serum and glucocorticoid induced protein kinase; SGK; PKBa1pha;
 KW 3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
 KW thrush; cancer; diabetes; obesity; antifungal; Candida infection; ss.

XX OS Rattus sp.

XX FH Key Location/Qualifiers
 FT CDS 1..1288

FT FT /*tag= a "serum and glucocorticoid induced protein
 FT /product= kinase (SGK)"
 FT /trans= except- (pos: 1, aa: Thr)
 FT /note= "the codon encoding Met at position 1 and
 FT Pro at position 318 are not given"

XX PN W0200036135-A2.

XX PD 22-JUN-2000.

XX PP 14-DEC-1999; 99WO-G504228.

XX PR 14-DEC-1998; 98US-0112114.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Thorner JW, Alessi DR, Torrance PD, Casamayor A;

XX DR WPI; 2000-442391/38.

XX DR P-PSDB; AAY93530.

XX FT Screening method identifying compounds which modulate protein kinase
 XX activity for use in treating fungal infections and cancer .

XX PS Disclosure; Fig 12; 155pp; English.

XX The present sequence encodes a rat serum and glucocorticoid induced
 CC protein kinase (SGK). The specification describes a screening method
 CC to identify a compound which modulates the activity of protein kinases
 CC from different sources, using host yeast cells. The method is used to
 CC identify a compound which modulates (inhibits) the activity of a
 CC protein kinase. Pkh1 or Pkh2 phosphorylate and activate Ypk1, Yrk2,
 CC SGK or protein kinase B-alpha (PKBa1pha). 3-phosphoinositide-dependent
 CC protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
 CC Yrk2 or SGK but not PKBa1pha or p70S6 kinase. Compounds identified by
 CC the methods are used to treat fungal infections e.g. thrush, and to
 CC treat cancer. To treat cancer, the compounds inhibit PKB, PDK1 or the

CC activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be
 CC used in the treatment of diabetes or obesity, and compounds which
 CC inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK
 CC (Ypk1 or Yrk2) can be used as an antifungal agent to treat Candida
 CC infections, e.g. thrush.

XX SQ Sequence 2365 BP; 591 A; 580 C; 518 G; 676 T; 0 other;

Alignment Scores:

Score: 1e-67 Length: 2365
 Pred. NO.: 582.50 Matches: 111
 Percent Similarity: 98.32% Conservative: 6
 Best Local Similarity: 93.28% Mismatches: 1
 Query Match: 93.35% Indels: 1
 DB: 21 Gaps: 1

US-10-000-039A-2_COPY_313_431 (1-119) x AAA46592 (1-2365)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 932 CCTCTCCAGCTGAAA---AATATCACCACCTCAGCAGGACCTGCTGGAGGGCTCCG 988
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
 DB 989 CAGAAGGACCGGACCAAGAGGCTGGGTGCCAAGGATGACTTTATGGAGATTAAAGATCAT 1048
 QY 41 ValPhePheSerLeuLeuAsnTyrAspAspLeuIleAsnLysLysIleThrProPhe 60
 DB 1049 ATTTTCTCTCTTTGATTAACTGGGATGATCTCATTAAATAGAAGATCAGCCGCCATTT 1108
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 DB 1109 AACCCAAATGTAGTGGGCCAGTGCACCTTCGGCAGCTTTGATCCGAGTTTACCGAGGAG 1168
 QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerVallys 100
 DB 1169 CCGTCCCGAGCTCCATCGCAGGTCCTGACAGCATCTTGTACAGCAGCATGTGAAA 1228
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 1229 GAAGCCGCGAAGCCCTTCTCGCTCTCTATGACACCTCCCTGTGTGATCTCTCTC 1285

RESULT 13

ABX42261

ID ABX42261 standard; cDNA; 404 BP.

XX AC ABX42261;

XX DT 20-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #7426.

XX KW Bovine; ss: EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX PN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-0960352.

XX PR 12-JAN-1999; 99US-115707P.

XX PR 11-JAN-2000; 2000US-0480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.
 XX PT New nucleic acid associated with lactation, and muscle and fat
 XX PT deposition, useful for genome mapping, gene identification and
 XX PT analysis, cattle breeding, or for genetically improving cattle
 XX PS Claim 2; SEQ ID NO 7436; 245pp; English.
 XX CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived
 CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC nucleic acid in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139.
 XX SQ Sequence 404 BP; 97 A; 108 C; 101 G; 98 T; 0 other;

Alignment Scores:
 Pred. No.: 2,36e-59 Length: 404
 Score: 512.00 Matches: 95
 Percent Similarity: 99.03% Conservative: 7
 Best Local Similarity: 92.23% Mismatches: 1
 Query Match: 82.05% Indels: 0
 DB: 25 Gaps: 0

US-10-000-039a-2_copy_313_431 (1-119) x ABX42261 (1-404)

QY 17 GluGluLeuLeuGlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGlu 36
 DB 13 GAAGCCCTCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 72
 QY 37 IleLysSerHisValPheSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLysLeu 56
 DB 73 APTAAGAATCATGCT 132
 QY 57 ThrProProPheAsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGlu 76
 DB 133 ACTCCCTCTTTTAAACCAAAATGTGAGCGGACCGACGACGACGACGACGACGACGAC 192
 QY 77 PheThrGluGluProValProAsnSerIleGlyLysSerProAspSerValLeuValThr 96
 DB 193 TTCCTAGAGAGCCCGGTCCTCCCACTCCATCCGCGCGGTCCTCCCTCTCTCTCTCA 252
 QY 97 AlaSerValLysGluAlaAlaGluAlaPheLeuGlyPheSerTyralaProThrAsp 116
 DB 253 GCACGCTCAAGAGAGCGGCTGAGGCGCTTCTCTGCGCTTCTCTGACCTCCATCCATG 312
 QY 117 SerPheLeu 119
 DB 313 TCTTCTCTC 321

RESULT 14

ABX38963
 ID ABX38963 standard; cDNA; 433 BP.
 XX AC ABX38963;
 XX DT 20-FEB-2003 (first entry)
 XX DE Bovine EST associated with lactation/muscle/fat deposition #4128.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX OS Bos Taurus.
 XX PN US2002137139-A1.
 XX PD 26-SEP-2002.
 XX PF 24-SEP-2001; 2001US-0960352.
 XX PR 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 DR WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX Claim 2; SEQ ID NO 4128; 245pp; English.
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived
 CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 15112 nucleotide
 CC sequences, appearing as ABX34836-ABX49947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a
 CC nucleic acid (comprising any of the 15112 nucleic acid sequences or its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 15112 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139.
 XX SQ Sequence 433 BP; 107 A; 115 C; 105 G; 106 T; 0 other;

Alignment Scores:
 Pred. No.: 1.46e-57 Length: 433
 Score: 499.00 Matches: 95

Percent Similarity: 98.08% Conservative: 7
 Best Local Similarity: 91.35% Mismatches: 1
 Query Match: 79.97% Indels: 1
 DB: 25 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x AEX38963 (1-433)

QY 17 GluCluLeuGluLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGlu 36
 DB 13 GAAGGCTCTCTCAGAGACAGGACAAAGAGGCTGGTGGCCAGGATGACTTTATGGAG 72
 QY 37 IleLysSerHisValPhePheSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLysLys 56
 DB 73 ATTAAGATCATGCT 132
 QY 57 ThrProPheAsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGlu 76
 DB 133 ACTCCCTCTTTAAACCAATGTGAGCGGACCGAGCTCGGACACTTTCATCTGAG 192
 QY 77 PheThrGluGluProValProAsnSerIleGlyLysSerProAspSerValLeuValThr 96
 DB 193 TTCCTGAGAGCGGCTCCCACTCCATCCGCGGCTCCCGAGAGCTCTCTCTCACA 252
 QY 97 AlaservValLysGluAlaGluAlaPheLeuGlyPheSerTyrAlaPro-ProThrAs 116
 DB 253 GCCAGGCTCAAGGAGCGGCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
 QY 116 pSerPheLeu 119
 DB 313 CTCTTTCTCTC 322

RESULT 15

AAH99535
 ID AAH99535 standard; CDNA; 1333 BP.

AC AAH99535;
 XX

DT 16-OCT-2001 (first entry)

DE Human protein encoding cDNA sequence SEQ ID NO:370.

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW anti-inflammation; anti-rheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; virucide;
 KW anti-HIV; fungicide; antitumor; cardiovascular; antianaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
 KW dermatological; antiasthmatic; antidiabetic; cyostatic;
 KW neuroprotective; antileptant; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.

OS Homo sapiens.

PN WO200153455-A2.

PD 26-JUL-2001.

XX 22-DEC-2000; 2000WO-US35017.

XX 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

XX (HYSE-) HYSEQ INC.

PA Tang YT, Liu C, Drmanac RT;

PI

DR

DR WP1; 2001-457603/49.

XX P-PSDB; AAM25594.

XX Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX Claim 1; Page 474; 1217pp; English.

XX

XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antitumor;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antiulcer; osteopathic; dermatological; antiasthmatic; antidiabetic;
 CC antidiabetic; cyostatic; neuroprotective; antileptant; antiparkinsonian;
 CC and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antitense therapy and vaccine
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 1333 BP; 315 A; 383 C; 349 G; 286 T; 0 other;

Alignment Scores:

Pred. No.: 3,48e-39 Length: 1333
 Score: 367.50 Matches: 72
 Percent Similarity: 75.63% Conservative: 18
 Best Local Similarity: 60.50% Mismatches: 26
 Query Match: 58.89% Indels: 3
 DB: 22 Gaps: 1

US-10-000-039A-2_COPY_313_431 (1-119) x AAH99535 (1-1333)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuGluGlyLeuLeu 20
 DB 898 CCCTACAGATCCCGGAGCGCCGACAGTGGCGGCTGTGACCTCTCTCAAGCTTCTC 957
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLysSerHis 40
 DB 958 CACAGGACCCAGGAGCGGCTGGCTCCAAAGCAGACCTTTCTTGAGATTAGAACCAT 1017
 QY 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
 DB 1018 GTATCTTCAGCCCATTAACCTGGATGACCTGTACCACAGAGGCTACTCCACATTC 1077
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 DB 1078 AACCAAAATGTGACAGGAGCTGCTGACTTGAGACATTTTGACCCAGAGTTCCACAGGAA 1137
 QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 DB 1138 GTGTGTCACAGTCCATTGGCTGTACCCCTGACACTGTG-----GCCAGCAGCTCT 1188
 QY 101 GluAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 1189 GGGGCTCAAGTGCATCTCTGGGATTTCTTTATGGCCAGAGGATGATGATCTTTG 1245

Search completed: August 11, 2003, 08:52:52

Job time : 449.545 secs

GeuCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 11, 2003, 08:36:05 ; Search time 3824.23 Seconds
(without alignments)
756.292 Million cell updates/sec

Title: US-10-000-039A-2_COPY_313_431

Perfect score: 624

Sequence: 1 PLQKPNITNSARHLEGLL.....KEAAEAFLEGSVAPPDSEFL 119

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=xlp
-q/cgn2_1/USPto_Spool/US10000039/runat_01082003_085419_27521/app_query.fasta_1.661
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039 -CGN_1_1_6635 -runat_01082003_085419_27521 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpi:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estow:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_plo:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	621	99.5	505	12	BM707970 UI-E-C11-
2	621	99.5	557	12	BM705333 UI-E-C11-
3	621	99.5	608	14	CB216184 NISC_nq02
4	621	99.5	630	14	CB216592 NISC_nq06
5	621	99.5	671	12	BC770190 602744954
6	621	99.5	721	12	RG778994 602666150
7	621	99.5	765	12	RG763506 602735901
8	621	99.5	847	13	BQ212589 AGENCOURT
9	621	99.5	859	13	BQ680126 AGENCOURT
10	621	99.5	873	10	BG756781 602710254
11	621	99.5	898	13	BQ687393 AGENCOURT
12	621	99.5	909	12	BQ006359 603615058
13	621	99.5	912	13	BQ676527 AGENCOURT
14	621	99.5	918	13	BQ838753 AGENCOURT
15	621	99.5	933	12	BG911625 602812771
16	621	99.5	969	12	B1333256 602996596
17	621	99.5	997	13	BQ691213 AGENCOURT
18	614	98.4	722	9	AV703365 AV703365
19	613	98.2	1068	13	B0151460 AGENCOURT
20	609	97.6	458	14	CB114428 K-EST0158
21	608	97.4	912	10	BE730437 601561461
22	606	97.1	561	14	CB154488 K-EST0212
23	605	97.0	899	12	BG769533 602744420
24	605	97.0	905	12	BG762465 602731909
25	601	96.3	537	10	BE285831 601096570
26	601	96.3	553	14	CD561854 B0438A08-
27	601	96.3	666	10	BE285821 601096554
28	601	96.3	726	14	CB954486 AGENCOURT
29	601	96.3	810	12	RG915496 602815720
30	601	96.3	858	10	BE287901 601096670
31	601	96.3	926	10	BF535805 602051316
32	601	96.3	957	12	B1553329 603236073
33	601	96.3	2435	11	AK086892 Mus muscu
34	597	95.7	456	9	A1265326 uk01q03.y
35	594	95.2	851	12	BT764996 603651130
36	590	94.6	682	10	BE307899 601096562
37	582	93.3	736	10	RG575325 602597986
38	581	93.1	547	10	RG712675 p411n.pk0
39	581	93.1	940	13	B0141033 60337256
40	570	91.3	790	12	B1662139 603305157
41	565	90.5	473	12	BJ071927 BJ071927
42	565	90.5	581	12	BJ040401 BJ040401
43	565	90.5	618	12	BJ040654 BJ040654
44	565	90.5	621	12	BJ063281 BJ063281
45	565	90.5	662	12	BJ094705 BJ094705

ALIGNMENTS

RESULT 1
BM707970
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

BM707970 505 bp mRNA linear EST 28-FEB-2002
UI-E-C11-aft-k-17-0-UI-rl UI-E-C11 Homo sapiens cDNA clone
UI-E-C11-aft-k-17-0-UI 5', mRNA sequence.
BM707970
BM707970.1 GI:19021228
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 505)

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548

CONTACT: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES
 source

1..505
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-aft-k-17-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-C11"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site:1: EcoR I; Site:2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

BASE COUNT 124 a 143 c 117 g 121 t
ORIGIN

Alignment Scores:
 Pred. No.: 1.03e-65 Length: 505
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 12 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119); x BM707970 (1-505)

QY 1 ProLeuGluProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 |||||||
Db 146 CCTCTCCAGCTGAACCAATATACAAATTCGCGACAGACACCTCTCGGAGGCGCTCGT 205
 |||||||

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
 |||||||
Db 206 CAGACGACAGACAAAGCGCTCGGCGCCCAAGGATGACTTCATGGAGATTAGAGTCAT 265
 |||||||

QY 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe 60
 |||||||
Db 266 GTCCTTCCTCTTAATTAACCTGGGATGATCTCATTAATAAGAGATTACTCCCTCTTT 325
 |||||||

QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 |||||||

Db 326 AACCAATGTGAGTGGCCCAAGACCTAGGCACCTTTGACCCCGAGTTTACGAAAGAG 385
 |||||||

QY 81 ProValProAsnSerIleGlyLysSerProAsnSerValLeuValThrAlaSerValLys 100
 |||||||

Db 386 CTTGTCCCAACTCCATTCGCAAGTCCCTCCGACAGGCTCCTGTCACAGCAGCGTCAAG 445
 |||||||

QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProPheThrAspSerPheLeu 119
 |||||||

Db 446 GAGCTGCCGAGGCTTCTTAGGCTTTTCTATGGCCCTCCACGAGCCTTTTCCTTC 502
 |||||||

RESULT 2
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BM705333 557 bp mRNA linear EST 28-FEB-2002
 UI-E-C11-agf-j-21-0-UI.r1 UI-E-C11 Homo sapiens CDNA clone
 UI-E-C11-agf-j-21-0-UI 5', mRNA sequence.
 BM705333
 BM705333.1 GI:19018591
 EST.
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 557)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548

COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES
 source

1..557
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-agf-j-21-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-C11"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site:1: EcoR I; Site:2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

BASE COUNT 133 a 148 c 130 g 144 t
ORIGIN

Alignment Scores:

Pred. No.: 1.16e-65 Length: 557
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 12 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x BM705333 (1-557)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 146 CCTCTCAGCTGAACCAATATACAAATTCCTCCAGACACTCTCTGGAGGCGCTCTG 205
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
 DB 206 CAGAAGCAGCAGGACAAAGCGCTCGGGGCCAAGGATGATCTCATGGAGATTAAAGATCAT 265
 QY 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProProPhe 60
 DB 266 GTCTTCCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 325
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 DB 326 AACCCAAATGTAGTGGGCCCAACGACCTACGCGCACTTTGACCCCGAGTTTACCGAAGAG 385
 QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 DB 386 CCTGTCCCACTCCATTCGCAAGTCNCTGACAGGCTCTCTGTCACAGCGGCTCAAG 445
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 446 GAAGCTCCGAGGCTTTCTTAGGCTTTTCTATGCGCTCCCGAGCTCTTTCTCTC 502

RESULT 3

LOCUS CB216184 608 bp mRNA linear EST 06-FEB-2003
 DEFINITION NISC_nq02c07.y1 NICHDS_HS_Ut2 Homo sapiens cDNA clone IMAGE:5937781
 5', mRNA sequence.
 ACCESSION CB216184
 VERSION CB216184.1 GI:28264376
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 CDNA Library Preparation: The I.M.A.G.E. Consortium/LLNL
 CDNA Library Arrayed by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Clone distribution: NCI-CCGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 info@image.llnl.gov
 Plate: LHAMJ3165 row: E column: 14
 Seq primer: M13Rpl reverse primer (AB1).
 Location/Qualifiers
 1. 608
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5937781"
 /sex="female"
 /tissue_type="normal endometrium, mid-secretory phase,
 cycle day 23"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NICHDS_HS_Ut2"
 /note="Organ: uterus; Vector: pCMV-SPORT6.1.ccdB (ResGen),

FEATURES

source
 1. 608
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5937781"
 /sex="female"
 /tissue_type="normal endometrium, mid-secretory phase,
 cycle day 23"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NICHDS_HS_Ut2"
 /note="Organ: uterus; Vector: pCMV-SPORT6.1.ccdB (ResGen),

Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
 Cloned unidirectionally from microquantity amounts of mRNA
 from normal endometrial tissue (mid-secretory phase, cycle
 day 23). Average insert size 1.6 kb. Library constructed
 by ResGen (Invitrogen Corporation)."

BASE COUNT 151 a 164 c 132 g 161 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.28e-65 Length: 608
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 14 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x CB216184 (1-608)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 59 CCTCTCAGCTGAACCAATATACAAATTCCTCCAGACACTCTCTGGAGGCGCTCTG 118
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLeuLysSerHis 40
 DB 119 CAGAAGCAGCAGGACAAAGCGCTCGGGGCCAAGGATGATCTCATGGAGATTAAAGATCAT 178
 QY 41 ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProProPhe 60
 DB 179 GTCTTCCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 238
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 DB 239 AACCCAAATGTAGTGGGCCCAACGACCTACGCGCACTTTGACCCCGAGTTTACCGAAGAG 298
 QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
 DB 299 CCTGTCCCACTCCATTCGCAAGTCNCTGACAGGCTCTCTGTCACAGCGGCTCAAG 358
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
 DB 359 GAAGCTCCGAGGCTTTCTTAGGCTTTTCTATGCGCTCCCGAGCTCTTTCTCTC 415

RESULT 4

LOCUS CB216592 630 bp mRNA linear EST 06-FEB-2003
 DEFINITION NISC_nq06g10.y1 NICHDS_HS_Ut2 Homo sapiens cDNA clone IMAGE:5938363
 5', mRNA sequence.
 ACCESSION CB216592
 VERSION CB216592.1 GI:28264784
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 CDNA Library Preparation: The I.M.A.G.E. Consortium/LLNL
 CDNA Library Arrayed by: National Institutes of Health Intramural
 Sequencing Center (NISC)
 Clone distribution: NCI-CCGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 info@image.llnl.gov
 Plate: LHAMJ3166 row: M column: 20
 Seq primer: M13Rpl reverse primer (AB1).
 Location/Qualifiers
 1. 630
 /organism="Homo sapiens"

FEATURES

source

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5938363"
/sex="female"
/tissue_type="normal endometrium, mid-secretory phase,
cycle day 23"
/lab_host="DH10B (Tl-resistant)"
/clone_lib="NICHDS_UT2"
/note="Organ: uterus; Vector: pCMV-Sport6.1.cdb (ResGen,
Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
Cloned unidirectionally from microquantity amounts of mRNA
from normal endometrial tissue (mid-secretory phase, cycle
day 23). Average insert size 1.6 kb. Library constructed
by ResGen (Invitrogen Corporation)."

BASE COUNT 153 a 167 c 141 g 169 t
ORIGIN
Alignment Scores:
Pred. No.: 1.34e-65 Length: 630
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 14 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x CB216592 (1-630)

Qy 1 ProLeuGlnLeuLysProAsnGluLeuArgHisLeuLeuGluGlyLeuLeu 20
Db 93 CCTCTCCAGCTGAACCAAAATATACAAATTCGCGAAGACACCTCTCGAGGGGCTCTG 152
Qy 21 GlnLysAspArgThrLysArgGluGlyAlaLysAspPheMetGluLeuLysSerHis 40
Db 153 CAGAGGACGACGACAAAGCGCTCGGGCCACGATGCTTCATGGAGATTAAAGATCAT 212
Qy 41 ValPhePheSerLeuLeuLeuTrpAspAspLeuLeuLeuLysLysLysLysLysLys 60
Db 213 GTCTTCTCTCTTAATTAATGAGTATGATCTCATTAATAAGAAAGATTACTCCCTCTT 272
Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db 273 AACCCAAATGTGAGTGGGCCCAACGACCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 332
Qy 81 ProValProAsnSerLeuGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 333 CCTGTCCCACTCCATTCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 392
Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 393 GAAGCTGCCGAGGCTTCTAGGCTTTCTAGGCTTTCTAGGCTTCTAGGCTTCTAGGCT 449

RESULT 5
BG770190 671 bp mRNA linear EST 15-MAY-2001
LOCUS 602744954F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4878085 5',
DEFINITION mRNA sequence.

ACCESSION BG770190
VERSION BG770190.1 GI:14080843
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 671)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

Tissue Procurement: ATCC/BCDB/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.N.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.N.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI763 row: k column: 14
High quality sequence slop: 669.

FEATURES

Location/Qualifiers
1..671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4878085"
/tissue_type="metanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_49"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. I"

BASE COUNT 160 a 171 c 152 g 188 t
ORIGIN

Alignment Scores:
Pred. No.: 1.45e-65 Length: 671
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 12 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x BG770190 (1-671)

Qy 1 ProLeuGlnLeuLysProAsnGluLeuArgHisLeuLeuGluGlyLeuLeu 20
Db 102 CCTCTCCAGCTGAACCAAAATATACAAATTCGCGAAGACACCTCTCGAGGGGCTCTG 161
Qy 21 GlnLysAspArgThrLysArgGluGlyAlaLysAspPheMetGluLeuLysSerHis 40
Db 162 CAGAGGACGACGACAAAGCGCTCGGGCCACGATGACTTCATGGAGATTAAAGATCAT 221
Qy 41 ValPhePheSerLeuLeuLeuTrpAspAspLeuLeuLysLysLysLysLysLys 60
Db 222 GTCTTCTCTCTTAATTAATGAGTATGATCTCATTAATAAGAAAGATTACTCCCTCTT 281
Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
Db 282 AACCCAAATGTGAGTGGGCCCAACGACCTACGCGACTTTGACCCCGAGTTTACCGAAGAG 341
Qy 81 ProValProAsnSerLeuGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 342 CCTGTCCCACTCCATTCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 401
Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 402 GAAGCTGCCGAGGCTTCTAGGCTTTCTAGGCTTTCTAGGCTTCTAGGCTTCTAGGCT 458

RESULT 6

BG778994 721 bp mRNA linear EST 15-MAY-2001
LOCUS 602666150F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4805724 5',
DEFINITION mRNA sequence.

ACCESSION BG778994
VERSION BG778994.1 GI:14049311
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 721)

QY 81 ProValProAsnSerIleGlySerProAspSerValLeuValThrAlaSerValLys 100
 DB 481 CCTGTCCCAACCTCCATGGCAAGTCCCTGACAGCGCTCTCGTTCACAGCGGTCACAG 540
 QY 101 GluAlaAlaGluAlaPheSerTyraLysSerValLeuValThrAlaSerValLys 119
 DB 541 GAAGTCCCGAGCGCTTCTTAGGCTTTCTATGCGCTCCACGCGACTCTTTCCTC 597

RESULT 8
 BQ212589 847 bp mRNA linear EST 02-MAY-2002
 LOCUS AGENCOURT_7675949 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095827
 DEFINITION 5', mRNA sequence.

ACCESSION BQ212589
 VERSION BQ212589.1 GI:20392956
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 847)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs.fda.gov

Tissue Procurement: ATCC/DCTD/DTP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L14M13368 row: n column: 20
 High quality sequence stop: 636
 Location/Qualifiers

FEATURES
 Source 1..847
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6095827"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_72"
 /note="Organ: skin; Vector: pCMV-SF006; Site: 1; NotI;
 Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 194 a 247 c 190 g 216 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,91e-65 Length: 847
 Score: 621.00 Matches: 118
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 13 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x BQ212589 (1-847)

QY 1 ProLeuGlnLeuLysProAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 264 CCTCTCCAGCTGAACCAATATACAAATATCCGACAGACCTCTCTGAGGCGCTCTG 323

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspSerPheMetGluLeuLysSerHis 40
 DB 324 CAGAAGGACAGGACAAAGCGCTCGGCGCAAGGATGACTTCATGAGATTACAGATCAT 383

QY 41 ValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysIleThrProPhe 60
 DB 384 GTCT 443

QY 61 AsnProAsnValSerGlyProAsnGlnLeuArgHisPheAspProGluPheThrGluLeu 80
 DB 444 AACCCAAATGTGAGTGGGCGCCAAACGACCTAGCGGACTTTGACCCGAGTTTACCGAAGAG 503

QY 81 ProValProAsnSerIleGlySerProAspSerValLeuValThrAlaSerValLys 100
 DB 504 CCTGTCCCAACCTCCATGGCAAGTCCCTGACAGCGCTCTCGTTCACAGCGGTCACAG 563

QY 101 GluAlaAlaGluAlaPheSerTyraLysSerValLeuValThrAlaSerValLys 119
 DB 564 GAAGTCCCGAGCGCTTCTTAGGCTTTCTATGCGCTCCACGCGACTCTTTCCTC 620

RESULT 9
 BQ680126 859 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT_8196937 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6265586
 DEFINITION 5', mRNA sequence.

ACCESSION BQ680126
 VERSION BQ680126.1 GI:21792805
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 859)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs.fda.gov

Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ruben Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L14M2427 row: 1 column: 19
 High quality sequence stop: 718
 Location/Qualifiers

FEATURES
 Source 1..859
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6263586"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_112"
 /note="Organ: skin; Vector: pOT87; Site: 1; XhoI; Site: 2;
 EcoRI; CDNA made by oligo-dr priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."

BASE COUNT 206 a 211 c 201 g 241 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,94e-65 Length: 859
 Score: 621.00 Matches: 118
 Percent Similarity: 100.0% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 13 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x BQ680126 (1-859)

QY 1 ProLeuGlnLeuLysProAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 DB 95 CCTCTCCAGCTGAACCAATATACAAATATCCGACAGACCTCTCTGAGGCGCTCTG 154

```

QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLleLysSerHis 40
DB 155 CAGAGGACAGGACAAAGCGCTCGGGCCAAAGGATGACTTTCATGAGATTAAAGATCAT 214
QY 41 ValPhePheSerLeuIleAsnTTPAspAspLeuIleAsnLysLysIleThrProProPhe 60
DB 215 GTCTCTCTCTCTTAATAACTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTT 274
QY 61 AsnProValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
DB 275 AACCCAAATGTAGTGGGCCCAAGCAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 334
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
DB 335 CCGTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 394
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
DB 395 GAAGCTGCCGAGGCTTTTCCTAGGCTTTTCCTATGCGCTCCACGAGACTCTTTCCTC 451

```

```

RESULT 10
BG756781
LOCUS
DEFINITION      873 bp mRNA linear EST 15-MAY-2001
                  mRNA sequence.
ACCESSION      BG756781
VERSION        BG756781.1 GI:14067434
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 873)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
                  cDNA Library Preparation: Ling Hong/Rubin Laboratory
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

```

```

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LMG1692 row: k column: 13
High quality sequence stop: 797.
Location/Qualifiers
1..873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4850820"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="NIH-MGC-48"
/clone_lib="NIH-MGC-48"
/notes="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH-MGC Library."

```

```

FEATURES
source
1..873

```

```

BASE COUNT      201 a 211 c 205 g 256 t
ORIGIN
Alignment Scores:
Pred. No.:      1,98e-65      Length:      873
Score:          621.00      Matches:      118
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.16%      Mismatches: 0

```

```

Query Match:      99.52%      Indels:      0
DB:              10          Gaps:        0
US-10-000-039a-2_copy_313_431 (1-119) x BG756781 (1-873)
QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluLleLysSerHis 20
DB 3 CCTCTCCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGACCTCTG 62
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLleLysSerHis 40
DB 63 CAGAGGACAGGACAAAGCGCTCGGGCCAAAGGATGACTTTCATGAGATTAAAGATCAT 122
QY 41 ValPhePheSerLeuIleAsnTTPAspAspLeuIleAsnLysLysIleThrProProPhe 60
DB 123 GTCTCTCTCTCTTAATAACTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTT 182
QY 61 AsnProValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
DB 183 AACCCAAATGTAGTGGGCCCAAGCAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAG 242
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
DB 243 CCGTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 302
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
DB 303 GAAGCTGCCGAGGCTTTTCCTAGGCTTTTCCTATGCGCTCCACGAGACTCTTTCCTC 359

```

```

RESULT 11
BG687393
LOCUS

```

```

DEFINITION      898 bp mRNA linear EST 15-JUL-2002
                  5', mRNA sequence.
ACCESSION      BG687393
VERSION        BG687393.1 GI:21812709
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE      1 (bases 1 to 898)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished
COMMENT        Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  cDNA Library Preparation: Rubin Laboratory
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LMG2364 row: j column: 03
                  High quality sequence stop: 574.
                  Location/Qualifiers
                  1..898
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6207842"
                  /tissue_type="ductal carcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH-MGC_110"
                  /note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
                  Site:2: EcoRI; cDNA made by oligo-dt priming.
                  Directionally cloned into EcoRI/XhoI sites using the
                  following 5' adaptor: GGCACGAG(G). Library constructed by
                  Ling Hong in the laboratory of Gerald M. Rubin (University
                  of California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies).
                  Note: this is a NIH-MGC Library."

```

```

FEATURES
source
1..898

```

us-10-000-039a-2_copy_313_431.rst

Tue Aug 12 10:55:57 2003

BASE COUNT 215 a 236 c 203 g 244 t
ORIGIN

Alignment Scores:
Pred. No.: 2,05e-65 Length: 898
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 13 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x B0687393 (1-898)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
|||||
Db 262 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTG 321
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
|||||
Db 322 CAGAAGGACAGACAAACGGCTCGGGCCCAAGATGATCTCATGAGATTAAGATGTCAT 381
QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProProPhe 60
|||||
Db 382 GTCTTCTCTCTTAATTAACCTGGATGATCTCATTAAGAAGATTAATCTCCCTCTTT 441
QY 61 AsnProAsnValSerGlyProAsnGlnLeuArgHisPheAspProGluPheThrGluGlu 80
|||||
Db 442 AACCCCAATGTGNGTGGGCCCAACGACCTACGGCNCITTGACCCCGAGTTTACCGAAGAG 501
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
|||||
Db 502 CTTGTCTCCCACTCCATTTGGCAAGTCCCTGACAGCGCTCTCTGTCACAGCCAGGTCAGAG 561
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 119
|||||
Db 562 GAAGCTGCGAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTCTC 618

RESULT 12

BM006359 603615058F1 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:5420871 5',
LOCUS mRNA sequence. EST 30-OCT-2001

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 909)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LCM1876 row: c column: 16

High quality sequence stop: 778.

FEATURES

source

1. 909
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5420871"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"

/note="Organ: pancreas; Vector: pOTB7; Site: 1; XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA Synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 229 a 233 c 214 g 233 t
ORIGIN

Alignment Scores:

Pred. No.: 2,08e-65 Length: 909
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 12 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x BM006359 (1-909)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
|||||
Db 264 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTG 323
QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
|||||
Db 324 CAGAAGGACAGACAAACGGCTCGGGCCCAAGATGATCTCATGAGATTAAGATGTCAT 383
QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProProPhe 60
|||||
Db 384 GTCTTCTCTCTTAATTAACCTGGATGATCTCATTAAGAAGATTAATCTCCCTCTTT 443
QY 61 AsnProAsnValSerGlyProAsnGlnLeuArgHisPheAspProGluPheThrGluGlu 80
|||||
Db 444 AACCCCAATGTGNGTGGGCCCAACGACCTACGGCNCITTGACCCCGAGTTTACCGAAGAG 503
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
|||||
Db 504 CTTGTCTCCCACTCCATTTGGCAAGTCCCTGACAGCGCTCTCTGTCACAGCCAGGTCAGAG 563
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 119
|||||
Db 564 GAAGCTGCGAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTCTC 620

RESULT 13

BM0676527

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 912)

NIH-MGC http://imgc.ncbi.nlm.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LCM2427 row: g column: 23

High quality sequence stop: 678.

FEATURES

Location/Qualifiers

source

1. 912
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6263470"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_112"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 218 a 227 c 213 g 254 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,096-65 Length: 912
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 13 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x B0676527 (1-912)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 |||||
 DB 95 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCGGAGGCTCTG 154
 |||||
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
 |||||
 DB 155 CAGAAGGACGAGGACAAACCGCTCGGGCCAGGATGATCTCATTAAGAAGATTAAAGATCAT 214
 |||||
 QY 41 ValPheSerLeuLeuLeuAsnTrpAspPheLeuLeuLeuLysIleThrProPhe 60
 |||||
 DB 215 GTCTTCTCTCTTAATTAACCTGGATGATCTCATTAAGAAGATTACTCCCTCTTT 274
 |||||
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 |||||
 DB 275 AACCCAAATGTAGTGGCCCAACACCTCGGACCTTGACCCGAGTTTACCCAGAG 334
 |||||
 QY 81 ProValProAsnSerIleGlyLysSerProAspPheValLeuValThrAlaSerValLys 100
 |||||
 DB 335 CCTGTCCCAACTCCCATTTGGCAAGTCCCTGACACGCTCTCGTCACAGCGGTCAAG 394
 |||||
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraProThrAspSerPheLeu 119
 |||||
 DB 395 GAAGCTCCGAGGCTTCTTAGGCTTTTCTATGCGCTCCACGAGCTCTTCTCTC 451
 |||||

RESULT 14

B0838753
 LOCUS
 DEFINITION B0838753 918 bp mRNA linear EST 16-OCT-2002
 AGENCOURT 8210094 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6258012
 5', mRNA sequence.

ACCESSION

B0838753

VERSION

B0838753.1 GI:24023148

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 918)

NIH-MGC http://img.ncbi.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: DCTD/DTP

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (I.M.I.)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LCM2413 row: d column: 13
 High quality sequence stop: 524.

FEATURES

Location/Qualifiers
 1..918
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6258012"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_112"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
 BASE COUNT 214 a 289 c 218 g 197 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,116-65 Length: 918
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 13 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x B0838753 (1-918)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
 |||||
 DB 95 CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCGGAGGCTCTG 154
 |||||
 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
 |||||
 DB 155 CAGAAGGACGAGGACAAACCGCTCGGGCCAGGATGATCTCATTAAGAAGATTAAAGATCAT 214
 |||||
 QY 41 ValPheSerLeuLeuLeuAsnTrpAspPheLeuLeuLeuLysIleThrProPhe 60
 |||||
 DB 215 GTCTTCTCTCTTAATTAACCTGGATGATCTCATTAAGAAGATTACTCCCTCTTT 274
 |||||
 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
 |||||
 DB 275 AACCCAAATGTAGTGGCCCAACACCTCGGACCTTGACCCGAGTTTACCCAGAG 334
 |||||
 QY 81 ProValProAsnSerIleGlyLysSerProAspPheValLeuValThrAlaSerValLys 100
 |||||
 DB 335 CCTGTCCCAACTCCCATTTGGCAAGTCCCTGACACGCTCTCGTCACAGCGGTCAAG 394
 |||||
 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyraProThrAspSerPheLeu 119
 |||||
 DB 395 GAAGCTCCGAGGCTTCTTAGGCTTTTCTATGCGCTCCACGAGCTCTTCTCTC 451
 |||||

RESULT 15

B0911625

LOCUS

B0911625

DEFINITION

602812771F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4944952

5', mRNA sequence.

ACCESSION

B0911625

VERSION

B0911625.1 GI:14292101

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 933)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LHAM0890 row: m column: 17
 High quality sequence stop: 799.
 Location/Qualifiers
 1..933
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone IMAGE:494952
 /tissue_type="anaplastic oligodendroglioma with lp/19q loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI_CGAP_Brn67"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
 BASE COUNT 226 a 250 c 215 g 242 t
 ORIGIN

Alignment Scores:
 pred No.: 2 15e-65 Length: 933
 Score: 621.00 Matches: 118
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.16% Mismatches: 0
 Query Match: 99.52% Indels: 0
 DB: 12 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x BG911625 (1-933)

QY	1	ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu	20
DB	204	CCTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCCTGGAGGGCCTCTG	263
QY	21	GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis	40
DB	264	CAGAAGGACAGACAAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGATCAT	323
QY	41	ValPhePheSerLeuIleAsnTrpAspAspLeuIleAsnLysLysIleThrProPhe	60
DB	324	GTCTTCTTCTCTTAATTAAGTGGGATGATCTCATTAATAAGAAGATTACTCCCTTTT	383
QY	61	AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu	80
DB	384	AAACCAAAATGTGAGTGGGCCCAACGACCTACGGCATTGACCCGAGTTTACCGAAGAG	443
QY	81	ProValProAsnSerIleGlySerProAspSerValLeuValThrAlaSerValLys	100
DB	444	CCTGTCCCAACTCCATTGGCAAGTCCCTCAGACGCTCCTCGTCACAGCCAGCGTCAAG	503
QY	101	GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu	119
DB	504	GAAGCTGGCGAGGCTTTCTTAGGCTTTTCTATGGCCCTCCACAGGACTCTTTCTCTC	560

Search completed: August 11, 2003, 11:27:43
 Job time: 3829.23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:36:34 ; Search time 113.591 Seconds
(without alignments)
462.401 Million cell updates/sec

Title: US-10-000-039A-2_COPY_313_431

Perfect score: 624
Sequence: 1 PLQKNVITNSARHLEGLL.....KENAEAFIGSFVAPPTDSPL 119

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Dgapop 6.0, Delxext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xl

-O=/cgn2.1/USPTO.spool/US10000039/runat_01082003_085419_27535/app_query.fasta_1.661
-DB=Issued_Patents.NA -QWFT=fastap -SUFFIX=trni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=13
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039_@CGN_1.145 -runat_01082003_085419_27535 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :

Issued_Patents.NA: *
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2.6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	624	100.0	2370	4	US-09-031-295-1
2	621	99.5	2311	2	US-08-712-709-6
3	621	99.5	2311	3	US-09-111-444-6
4	621	99.5	2311	3	US-09-541-228-6
5	621	99.5	2311	4	US-09-016-434-772
6	203.5	32.6	257	1	US-08-700-575-4
7	187	30.0	1607	2	US-08-749-902-4
8	187	30.0	1607	4	US-09-016-434-126
9	184	29.5	1732	4	US-09-430-564-1
10	184	29.5	2146	4	US-09-842-307-1
11	181.5	29.1	387	3	US-09-474-922A-2
12	175	28.0	2196	1	US-08-313-274-1

Sequence 1, Appli
Sequence 3, Appli
Sequence 70, Appli
Sequence 138, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1471, Ap
Sequence 72, Appli
Sequence 24, Appli
Sequence 25, Appli
Sequence 24, Appli
Patent No. 5266464
Sequence 3, Appli
Sequence 143, Ap
Sequence 25, Appli
Sequence 25, Appli
Sequence 25, Appli
Sequence 39, Appli
Sequence 68, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 35, Appli
Sequence 14, Appli
Sequence 3, Appli
Sequence 136, App
Sequence 48, Appli
Sequence 134, App
Sequence 61, Appli
Sequence 61, Appli
Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-09-031-295-1
; Sequence 1, Application US/09031295
; Patent No. 6326181

GENERAL INFORMATION:

APPLICANT: LANG, Florian
APPLICANT: WALDEGER, Tubingen
TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LAHDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,295
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 197-08-173.8

FILING DATE: 28-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sandercock, Colin G.

REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 058315/0123

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 2370 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 43..1335
US-09-031-295-1

Alignment Scores:

Pred. No.: 7,41e-74 Length: 2370
Score: 624.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x US-09-031-295-1 (1-2370)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
DB 979 CCTCTCCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGGAGGCTCCTG 1038
QY 21 GlnLysAspArgThrLysArgLeuGluAlaLysAspPheMetGluLysSerHis 40
DB 1039 CAGAGGACAGACAAACGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTCAT 1098
QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
DB 1099 GTCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTT 1158
QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
DB 1159 AACCCAAATGTGAGTGGGCCCAAGCAGCTACGGCCTTGACCCGAGTTTACCGAAGAG 1218
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
DB 1219 CTGTGCCCACTCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 1278
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
DB 1279 GAAGTCCCGAGGCTTCTTAGGCTTTCTCTATGCGCTCCACCGGACTCTTCTCCTC 1335

RESULT 2

US-08-712-709-6

; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-08-712-709-6

Alignment Scores:
Pred. No.: 1.81e-73 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 2 Gaps: 0

US-10-000-039A-2_COPY_313_431 (1-119) x US-08-712-709-6 (1-2311)

QY 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
DB 946 CCTCTCCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGGAGGCTCCTG 1005
QY 21 GlnLysAspArgThrLysArgLeuGluAlaLysAspPheMetGluLysSerHis 40
DB 1006 CAGAGGACAGACAAACGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTCAT 1065
QY 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
DB 1066 GTCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTT 1125
QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
DB 1126 AACCCAAATGTGAGTGGGCCCAAGCAGCTACGGCCTTGACCCGAGTTTACCGAAGAG 1185
QY 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
DB 1186 CTGTGCCCACTCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 1245
QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
DB 1246 GAAGTCCCGAGGCTTCTTAGGCTTTCTCTATGCGCTCCACCGGACTCTTCTCCTC 1302

RESULT 3

US-09-111-444-6

; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 1.5

APPLICATION NUMBER: US/09/111,444

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 2311 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY:

CLONE: Consensus

US-09-111-444-6

Alignment Scores:
Pred. No.: 1,81e-73 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservativity: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 3 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x US-09-111-444-6 (1-2311)

Qy 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
Db 946 CCTCTCCAGCTGAACCAATATACAAATTCGCAAGACACCTCTCGAGGGCTCTCTG 1005
Qy 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db 1006 CAGAGGACAGGACAGGCGCTCGGGCCAGGATGACTTCATGGAGATTAGAGTCAT 1065
Qy 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
Db 1066 GTCT 1125
Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
Db 1126 AACCAATGTAGTGGGCCCCAGGACCTAGCGCACTTTGACCCCGAGTTTACCGAGAG 1185
Qy 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 1186 CCTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTCTACAGCCAGCGTCAAG 1245
Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 1246 GAAGCTGCCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCT 1302

RESULT 4

US-09-541-228-6

Sequence 6, Application US/09541228

Patent No. 6232077

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Guegler, Karl J.

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/712,709

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0118 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 2311 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

LIBRARY:

CLONE: Consensus

US-09-541-228-6

Alignment Scores:
Pred. No.: 1,81e-73 Length: 2311
Score: 621.00 Matches: 118
Percent Similarity: 100.00% Conservativity: 1
Best Local Similarity: 99.16% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 3 Gaps: 0

US-10-000-039a-2_COPY_313_431 (1-119) x US-09-541-228-6 (1-2311)

Qy 1 ProLeuGlnLeuLysProAsnIleThrAsnSerAlaArgHisLeuLeuGluGlyLeuLeu 20
Db 946 CCTCTCCAGCTGAACCAATATACAAATTCGCAAGACACCTCTCGAGGGCTCTCTG 1005
Qy 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluIleLysSerHis 40
Db 1006 CAGAGGACAGGACAGGCGCTCGGGCCAGGATGACTTCATGGAGATTAGAGTCAT 1065
Qy 41 ValPhePheSerLeuLeuAsnTrpAspLeuLeuAsnLysLysIleThrProPhe 60
Db 1066 GTCT 1125
Qy 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 80
Db 1126 AACCAATGTAGTGGGCCCCAGGACCTAGCGCACTTTGACCCCGAGTTTACCGAGAG 1185
Qy 81 ProValProAsnSerIleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
Db 1186 CCTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTCTACAGCCAGCGTCAAG 1245
Qy 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 119
Db 1246 GAAGCTGCCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCT 1302

RESULT 5

US-09-016-434-772

Sequence 772, Application US/09016434

Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

1 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
2 TITLE OF INVENTION: PATHWAY GENE EXPRESSION
3 NUMBER OF SEQUENCES: 1490
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
6 STREET: 3174 PORTER DRIVE
7 CITY: PALO ALTO
8 STATE: CALIFORNIA
9 COUNTRY: USA
10 ZIP: 94304
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
16 CURRENT APPLICATION DATA:
17 APPLICATION NUMBER: US/09/016,434
18 FILING DATE: HEREWITH
19 CLASSIFICATION:
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER:
22 FILING DATE:
23 CLASSIFICATION:
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Zeller, Karen J.
26 REGISTRATION NUMBER: 37,071
27 REFERENCE/DOCKET NUMBER: PA-0002 US
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (650) 855-0555
30 TELEFAX: (650) 845-4166
31 INFORMATION FOR SEQ ID NO: 772:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 2311 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37 IMMEDIATE SOURCE:
38 LIBRARY: MMLRZDT01
39 CLONE: 477245
40 US-09-016-434-772
41
42 Alignment Scores:
43 Pred. No.: 1,81e-73 Length: 2311
44 Score: 621.00 Matches: 118
45 Percent Similarity: 100.00% Conservative: 1
46 Best Local Similarity: 99.16% Mismatches: 0
47 Query Match: 99.52% Indels: 0
48 DB: 4 Gaps: 0
49
50 US-10-000-039a-2_COPY_313_431 (1-119) x US-09-016-434-772 (1-2311)
51 QY 1 ProLeuGlnLeuLysProAsnLleThrAsnSerAlaArgHisLeuLeuGluGlyLeu 20
52 946 CCTCTCCAGCTGAGCAACCAATTTACAAATTCGCCAAGACACCTCTCGGAGGCCCTCG 1005
53 QY 21 GlnLysAspArgThrLysArgLeuGlyAlaLysAspPheMetGluLleLysSerHis 40
54 1006 CAGAAGGACAGACAAAGCGCTCGGGCCCAAGCATGCTTCATGGAGATTAAAGATCAT 1065
55 QY 41 ValPhePheSerLeuLeuAsnTTPAspAspLeuLeuLeuLysLysLleThrProPhe 60
56 1066 GTCTTCTCTCTTAATTAAGTGGGATGATCTCATTAATTAAGAGATTACTCCCTCTTT 1125
57 QY 61 AsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGluGlu 80
58 1126 AACCAATATGAGTGGGCCCAAGACCATCTAGGACATTTGACCCCGAGTTTACCGAAG 1185
59 QY 81 ProValProAsnSerLleGlyLysSerProAspSerValLeuValThrAlaSerValLys 100
60 1186 COTGTCCCAACTCCATTGGCAAGTCCCTCGACAGCGTCTCGTCACAGCCAGCGTCAAG 1245
61 QY 101 GluAlaAlaGluAlaPheLeuGlyPheSerTyrAlaProThrAspSerPheLeu 119
62 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

1246 GAAGTCCGCGAGGCTTTCTCTAGGCTTTCTCTATGCGCTCCACAGGACTCTTTCTC 1302
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000


```

; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 2
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-474-922A-2

```

Alignment Scores:		
Pred. No.:	1-41e-15	387
Score:	181.50	35
Percent Similarity:	63.16%	13
Best Local Similarity:	46.05%	27
Query Match:	29.09%	1
DB:	3	1
US-10-000-039A-2 COPY	313 431 (1-119)	x US-09-474-922A-2 (1-387)

US-10-000-039a-2_COPY_313_431 (1-119) x US-09-256-465-1 (1-1599)

QY 1 ProLeuGlnLeuLysProAsnLleThrAsnSerAlaAArgHisLeuLeuGluGlyLeuLeu 20
 Db 1195 CCGCGCAGCTCAGCCCC-----GAGGCCAAGTCCCTGCTTGTGGGTGCTT 1242
 QY 21 GlnLysAspArgThrLysArgLeu---GlyAlaLysAspPheMetGluLeuLysSer 39
 Db 1243 AAGAGGACCCCAAGCAGAGGCTTGTGGGGGCCCGAGGATCCCAAGGAGTCAATGAG 1302
 QY 40 HisValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysLleThrProPro 59
 Db 1303 CACAGGTCTCTCTCAGCATCAACTGGCAGGAGTGTCCAGAGAAGTCTCTGCACCC 1362
 QY 60 PheAsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 79
 Db 1363 TTCAACCTCAGGTCCAGTCCGAGGTCGACACAGGTACTTCGATGATGAATTTACCGCC 1422
 QY 80 GluProVal 82
 Db 1423 CAGTCCATC 1431

RESULT 14

US-09-167-322-3
 ; Sequence 3, Application US/09167322
 ; Patent No. 6365151

GENERAL INFORMATION:

APPLICANT: Allegheny University of the Health
 Sciences, Harpeth, Michael S.
 England, James M.

TITLE OF INVENTION: CANCER VACCINE

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seidel, Gonda, Lavoigna & Monaco, P.C.

STREET: Suite 1800, Two Penn Center Plaza

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19102

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/167,322

FILING DATE: 07-Oct-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/00582

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Monaco, Daniel A.

REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 7933-33 PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-8383

TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1599 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-167-322-3

Alignment Scores:

Pred. No.: 1,31e-13 Length: 1599
 Score: 173.50 Matches: 36
 Percent Similarity: 57.83% Conservative: 12
 Best Local Similarity: 43.37% Mismatches: 30
 Query Match: 27.80% Indels: 5

DB: 4 Gaps: 2

US-10-000-039a-2_COPY_313_431 (1-119) x US-09-167-322-3 (1-1599)

QY 1 ProLeuGlnLeuLysProAsnLleThrAsnSerAlaAArgHisLeuLeuGluGlyLeuLeu 20
 Db 1195 CCGCGCAGCTCAGCCCC-----GAGGCCAAGTCCCTGCTTGTGGGTGCTT 1242
 QY 21 GlnLysAspArgThrLysArgLeu---GlyAlaLysAspPheMetGluLeuLysSer 39
 Db 1243 AAGAGGACCCCAAGCAGAGGCTTGTGGGGGCCCGAGGATCCCAAGGAGTCAATGAG 1302
 QY 40 HisValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysLleThrProPro 59
 Db 1303 CACAGGTCTCTCTCAGCATCAACTGGCAGGAGTGTCCAGAGAAGTCTCTGCACCC 1362
 QY 60 PheAsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 79
 Db 1363 TTCAACCTCAGGTCCAGTCCGAGGTCGACACAGGTACTTCGATGATGAATTTACCGCC 1422
 QY 80 GluProVal 82
 Db 1423 CAGTCCATC 1431

RESULT 15

US-09-417-197-70

; Sequence 70, Application US/09417197

; Patent No. 6518021

GENERAL INFORMATION:

APPLICANT: Ole THASTRUP, et al.

TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To A

FILE REFERENCE: 3759-0110P

CURRENT APPLICATION NUMBER: US/09/417,197

CURRENT FILING DATE: 1999-10-07

NUMBER OF SEQ ID NOS: 143

SOFTWARE: Patent In version 3.0

SEQ ID NO 70

LENGTH: 2181

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: PKB-BGFP fusion

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(2178)

US-09-417-197-70

Alignment Scores:

Pred. No.: 8.3e-13 Length: 2181
 Score: 169.00 Matches: 45
 Percent Similarity: 50.00% Conservative: 16
 Best Local Similarity: 36.89% Mismatches: 47
 Query Match: 27.08% Indels: 14
 Gaps: 4

US-10-000-039a-2_COPY_313_431 (1-119) x US-09-417-197-70 (1-2181)

QY 1 ProLeuGlnLeuLysProAsnLleThrAsnSerAlaAArgHisLeuLeuGluGlyLeuLeu 20
 Db 1105 CCGCGCAGCTTGTGCTCC-----GAGGCCAAGTCCCTGCTTTCAGGGGTGCTC 1152
 QY 21 GlnLysAspArgThrLysArgLeu---GlyAlaLysAspPheMetGluLeuLysSer 39
 Db 1153 AAGAGGACCCCAAGCAGAGGCTTGTGGGGGCCCGAGGATCCCAAGGAGTCAATGAG 1212
 QY 40 HisValPhePheSerLeuLeuAsnTrpAspAspLeuLeuAsnLysLysLleThrProPro 59
 Db 1213 CAGTCTCTCTTTCGGGTATCGTGTGGCAGCAGTGTACGACAGAGTCTAGCCACCC 1272
 QY 60 PheAsnProAsnValSerGlyProAsnGluLeuArgHisPheAspProGluPheThrGlu 79
 Db 1273 TTCAAGCCCCCAGGTTCAGTCCGAGACTGACACACAGGTATTTTTCATGAGGAGTTCACGGCC 1332

Search completed: August 11, 2003, 11:30:24
Job time : 118.591 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 12:13:37 : Search time 2015 seconds
(without alignments)
10171.575 Million cell updates/sec

Title: US-10-000-039a-1_copy_980_1480
Perfect score: 501
Sequence: 1 CTCCTCAGCTGAACCAAT.....TAGCAATCTTATTCACACT 501

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	501	100.0	2370	6	AR179441	AR179441 Sequence
2	501	100.0	2370	6	AX002570	AX002570 Sequence
3	501	100.0	2370	6	AX337834	AX337834 Sequence
4	501	100.0	2370	6	AX411211	AX411211 Sequence
5	501	100.0	2370	9	HSRNAS1PK	Y10032 H. sapiens m
6	499.4	99.7	2281	6	AX017284	AX017284 Sequence
7	499.4	99.7	2281	6	AX524988	AX524988 Sequence
8	499.4	99.7	2281	6	BD134455	BD134455 Human nuc
9	499.4	99.7	2311	6	AR151390	AR151390 Sequence
10	499.4	99.7	2311	6	AR270209	AR270209 Sequence
11	499.4	99.7	2311	6	BD021928	BD021928 Human pro
12	499.4	99.7	2354	6	AX553549	AX553549 Sequence
13	499.4	99.7	2364	9	AK098509	AK098509 Homo sapi
14	499.4	99.7	2382	9	AF153609	AF153609 Homo sapi
15	499.4	99.7	2382	9	BC001263	BC001263 Homo sapi
16	499.4	99.7	3196	6	AX713538	AX713538 Sequence
17	499.4	99.7	3196	9	AK055077	AK055077 Homo sapi
18	365.6	73.0	1920	4	AF139639	AF139639 Oryctolag
19	359	71.7	1296	6	AX056374	AX056374 Sequence
20	315.2	62.9	2429	10	AF139638	AF139638 Mus muscu
21	314.2	62.7	1659	10	BC002222	BC002222 Mus muscu
22	314.2	62.7	2423	10	BC005720	BC005720 Mus muscu
23	314.2	62.7	2426	10	AF205855	AF205855 Mus muscu
24	314.2	62.7	2432	6	AX056375	AX056375 Sequence
25	311	62.1	5718	9	HSSGK	AX056375 Sequence
26	309.4	61.8	113673	9	AL138839	AL138839 Sequence
27	299.4	59.8	2435	10	RATSGPK	AL138839 Sequence
28	275.2	54.9	216056	2	AC103256	AL138839 Sequence
29	254	50.7	694	6	AX201922	AL138839 Sequence
30	244	48.7	183140	10	AC124486	AL138839 Sequence
31	232.6	46.4	1417	5	AF057138	AL138839 Sequence
32	207.2	41.4	2339	5	BC052134	AL138839 Sequence
33	205	40.9	2470	5	SAC223715	AL138839 Sequence
34	205	40.9	3105	5	SAC223716	AL138839 Sequence
35	157.2	31.4	142609	2	AC118095	AL138839 Sequence
36	157.2	31.4	240716	2	AC112280	AL138839 Sequence
37	155.8	31.1	220350	2	AC107830	AL138839 Sequence
38	151.4	30.2	208405	2	AC114405	AL138839 Sequence
39	140.6	28.1	431	6	AX305574	AL138839 Sequence
40	132.4	26.4	1346	6	AX056376	AL138839 Sequence
41	132.4	26.4	1812	6	AX056472	AL138839 Sequence
42	132.4	26.4	1812	6	AX399684	AL138839 Sequence
43	132.4	26.4	1812	9	AF169034	AL138839 Sequence
44	132.4	26.4	1815	9	BC014037	AL138839 Sequence
45	132.4	26.4	1834	6	AX026166	AL138839 Sequence

ALIGNMENTS

RESULT 1
AR179441
LOCUS AR179441
DEFINITION Sequence 1 from patent US 6326181.
ACCESSION AR179441
VERSION AR179441.1 GI:20220996
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2370)
AUTHORS Lang, F. and Waldegger, S.
TITLE Cell volume-regulated human kinase h-sgk
JOURNAL Patent: US 6326181-A 1 04-DEC-2001;
FEATURES Location/Qualifiers

Linear PAT 20-APR-2002
2370 bp
DNA


```
source
1..2370
/organism="unknown"
BASE COUNT      636 a      517 c      513 g      704 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 501; DB 6; Length 2370;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCTCCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTCGGAGGCGCTCTCC 60
DB 980 CTCTCCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTCGGAGGCGCTCTCC 1039
QY 61 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 120
DB 1040 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 1099
QY 121 TCTTCTTCCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTTA 180
DB 1100 TCTTCTTCCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTTA 1159
QY 181 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGACCTTTGACCCCGAGTTTACCGAAGGC 240
DB 1160 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGACCTTTGACCCCGAGTTTACCGAAGGC 1219
QY 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCCTCGTCACAGCGAGCTCAAGG 300
DB 1220 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCCTCGTCACAGCGAGCTCAAGG 1279
QY 301 AGCTGCCAGGCTTTCCTAGGCTTTTCTATGCGCTCCCGAGGACTCTTCTCTGAA 360
DB 1280 AGCTGCCAGGCTTTCCTAGGCTTTTCTATGCGCTCCCGAGGACTCTTCTCTGAA 1339
QY 361 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTGTAGTACCT 420
DB 1340 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTGTAGTACCT 1399
QY 421 TTGTGGAGCGCCAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 480
DB 1400 TTGTGGAGCGCCAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 1459
QY 481 TTAGCAATCTTATGCACT 501
DB 1460 TTAGCAATCTTATGCACT 1480

RESULT 3
AX337834
LOCUS      AX337834
DEFINITION Sequence 8343 from Patent WO0194629.
ACCESSION  AX337834
VERSION     AX337834.1 GI:18128553
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrikan,S., Soppet,D.R. and Weaver,Z.
TITLE       Cancer gene determination and therapeutic screening using signature
            gene sets
JOURNAL     Patent: WO 0194629-A 8343 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES
source
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT      636 a      517 c      513 g      704 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 501; DB 6; Length 2370;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCTCCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTCGGAGGCGCTCTCC 60
DB 980 CTCTCCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTCGGAGGCGCTCTCC 1039
QY 61 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 120
DB 1040 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 1099
QY 121 TCTTCTTCCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTTA 180
DB 1100 TCTTCTTCCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTACTCCCTTTTA 1159
QY 181 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGACCTTTGACCCCGAGTTTACCGAAGGC 240
DB 1160 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGACCTTTGACCCCGAGTTTACCGAAGGC 1219
QY 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCCTCGTCACAGCGAGCTCAAGG 300
DB 1220 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCCTCGTCACAGCGAGCTCAAGG 1279
QY 301 AGCTGCCAGGCTTTCCTAGGCTTTTCTATGCGCTCCCGAGGACTCTTCTCTGAA 360
DB 1280 AGCTGCCAGGCTTTCCTAGGCTTTTCTATGCGCTCCCGAGGACTCTTCTCTGAA 1339
QY 361 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTGTAGTACCT 420
DB 1340 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTGTAGTACCT 1399
QY 421 TTGTGGAGCGCCAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 480
DB 1400 TTGTGGAGCGCCAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 1459
QY 481 TTAGCAATCTTATGCACT 501
DB 1460 TTAGCAATCTTATGCACT 1480

RESULT 2
AX002570
LOCUS      AX002570
DEFINITION Sequence 1 from Patent EP0861896.
ACCESSION  AX002570
VERSION     AX002570.1 GI:7242111
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2370)
AUTHORS     Lang,F.P. and Waldegger,S.D.
TITLE       Cell volume regulated human kinase h-sgk
JOURNAL     Patent: EP 0861896-A 1 02-SEP-1998;
            DADE BEHRING MARBURG GMBH (DE)
FEATURES
source
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/strain="CAUCASIAN"
/db_xref="taxon:9606"
/clone="HSRNASPK"
BASE COUNT      636 a      517 c      513 g      704 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Score 501; DB 6; Length 2370;
```

QY 61 AGAAGACAGGACAAAGCGGCTCGGGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 120
|||||
Db 1040 AGAAGACAGGACAAAGCGGCTCGGGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 1099
QY 121 TCTCTCTCTCTTAATTAAGTGGATGATCTCAATTAATAGAGATTAATTAAGATTAATTA 180
|||||
Db 1100 TCTCTCTCTCTTAATTAAGTGGATGATCTCAATTAATAGAGATTAATTAAGATTAATTA 1159
QY 181 ACCCAATGTGAGTGGGCGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 240
|||||
Db 1160 ACCCAATGTGAGTGGGCGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 1219
QY 241 CTGTCCCAACTCCATTTGGAAGTCCCTGACAGGCTCTGCTGACAGGCGGTCAAGG 300
|||||
Db 1220 CTGTCCCAACTCCATTTGGAAGTCCCTGACAGGCTCTGCTGACAGGCGGTCAAGG 1279
QY 301 AAGCTGCGGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGGACTCTTCTCTGAA 360
|||||
Db 1280 AAGCTGCGGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGGACTCTTCTCTGAA 1339
QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATAGTACCT 420
|||||
Db 1340 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATAGTACCT 1399
QY 421 TTTGGTGGAGCGCCAGCTGACAGGATCTTACAAGAGATTGACATCTCTGGAAGC 480
|||||
Db 1400 TTTGGTGGAGCGCCAGCTGACAGGATCTTACAAGAGATTGACATCTCTGGAAGC 1459
QY 481 TTAGCAATCTTATTGCACACT 501
|||||
Db 1460 TTAGCAATCTTATTGCACACT 1480

RESULT 4

AX411211
LOCUS AX411211 2370 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3858 from Patent WO229103.
ACCESSION AX411211
VERSION AX411211.1 GI:21443916
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Alvarres,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3858 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source
1. .2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. Y10032"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN

Query Match 100.0%; Score 501; DB 6; Length 2370;
Best Local Similarity 100.0%; Pred. No. 3.4e-137;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATTACAAATTCGCCAAGACACCTCTGAGGGGCTCTCTGC 60
|||||
Db 980 CTCTCCAGCTGAACCAATATTACAAATTCGCCAAGACACCTCTGAGGGGCTCTCTGC 1039
QY 61 AGAAGACAGGACAAAGCGGCTCGGGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 120
|||||
Db 1040 AGAAGACAGGACAAAGCGGCTCGGGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 1099
QY 121 TCTCTCTCTCTTAATTAAGTGGATGATCTCAATTAATAGAGATTAATTAAGATTAATTA 180
|||||

Db 1100 TCTCTCTCTCTTAATTAAGTGGATGATCTCAATTAATAGAGATTAATTAAGATTAATTA 1159
QY 181 ACCCAATGTGAGTGGGCGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 240
|||||
Db 1160 ACCCAATGTGAGTGGGCGCCCAAGGATGACATTCATGGAGATTAAAGATCATG 1219
QY 241 CTGTCCCAACTCCATTTGGAAGTCCCTGACAGGCTCTGCTGACAGGCGGTCAAGG 300
|||||
Db 1220 CTGTCCCAACTCCATTTGGAAGTCCCTGACAGGCTCTGCTGACAGGCGGTCAAGG 1279
QY 301 AAGCTGCGGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGGACTCTTCTCTGAA 360
|||||
Db 1280 AAGCTGCGGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGGACTCTTCTCTGAA 1339
QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATAGTACCT 420
|||||
Db 1340 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATAGTACCT 1399
QY 421 TTTGGTGGAGCGCCAGCTGACAGGATCTTACAAGAGATTGACATCTCTGGAAGC 480
|||||
Db 1400 TTTGGTGGAGCGCCAGCTGACAGGATCTTACAAGAGATTGACATCTCTGGAAGC 1459
QY 481 TTAGCAATCTTATTGCACACT 501
|||||
Db 1460 TTAGCAATCTTATTGCACACT 1480

RESULT 5

HSRNASTPK
LOCUS HSRNASTPK 2370 bp mRNA linear PRI 16-MAY-1997
DEFINITION H.sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION Y10032
VERSION Y10032.1 GI:1834510
KEYWORDS serine/threonine protein kinase; SGK gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Waldegger,S., Barth,P., Raber,G. and Lang,F.
TITLE Cloning and characterization of a putative human serine/threonine
JOURNAL isoenzyme alterations of cell volume
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 94 (9), 4440-4445 (1997)
PUBMED 97272242
REFERENCE 2 (bases 1 to 2370)
AUTHORS Waldegger,S.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1996) S. Waldegger, University of Tuebingen,
Physiology I, Gmelinstr. 5, D-72076 Tuebingen, FRG

FEATURES

source
1. .2370
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
1. .2370
/gene="sgk"
43. .1338
/gene="sgk"
/codon_start=1
/product="serine/threonine protein kinase"
/protein_id="CAA71138.1"
/db_xref="GI:1834511"
/translation="MTVKTEAAKGLTYSRMGVYAILIAFMKORMGLNDFIQKIAN
NSYACKHPEVQSILKISQPELNMNAPSPSPSQINLGPSNHPKSDHFLLK
VIGSGFRVLLARHKAEEFYAVKVLQKRALKKKEKHINSEKLVLLKNVHPFLV
GLHFSFGADKLYFVLDYINGGELFYHLQRCFLPEPRARYAARIALGALYHSLNI
VYRDLKPNLIDSGHIVLTDFTCKENIEHNSTSTFCGTPPEYLAPEVLHKOPYDR
TVDMCLGAVLYEMLYLPPPYSRNTAEWYDNLNKLQLENITNSARHLGLQK
DRTKLGAQDDFWEIKSHVFTSLINWDDLINKKITPPPNVNSGPNELRHPDPETEE


```

QY 61 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 120
Db 940 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 999
QY 121 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
Db 1000 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1059
QY 181 ACCCAATGTAGTGGGGCCCAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 240
Db 1060 ACCCAATGTAGTGGGGCCCAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1119
QY 241 CTGTGCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 300
Db 1120 CTGTGCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 1179
QY 301 AAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCGCTCCCAAGGATGAGGATGAGGATGAGG 360
Db 1180 AAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCGCTCCCAAGGATGAGGATGAGGATGAGG 1239
QY 361 CCCTGTTAGGCTTGGTTTAAAGATTTTATGTTGTTTTCGAATGTTTATGTTAGGCT 420
Db 1240 CCCTGTTAGGCTTGGTTTAAAGATTTTATGTTGTTTTCGAATGTTTATGTTAGGCT 1299
QY 421 TTTGGTGGAGCGGCGGAGCTGACAGGATCTTACAAGAGAAATTTGACATCTCTGGAAGC 480
Db 1300 TTTGGTGGAGCGGCGGAGCTGACAGGATCTTACAAGAGAAATTTGACATCTCTGGAAGC 1359
QY 481 TTAGCAATCTTATTGCACACT 501
Db 1360 TTAGCAATCTTATTGCACACT 1380

```

```

RESULT 8
BD134455
LOCUS Human nucleic acid sequence originating in mammary tumor tissue.
DEFINITION
ACCESSION BD134455
VERSION BD134455.1 GI:23229400
KEYWORDS JP 2002506643-A/33.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2281)
Spelt,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
Rosenthal,A.
Human nucleic acid sequence originating in mammary tumor tissue
Patent: JP 2002506643-A 33 05-MAR-2002.
METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
OS Homo sapiens (human)
PN JP 2002506643-A/33
PD 05-MAR-2002
PF 19-MAR-1999 JP 2000536852
PR 20-MAR-1998 DE 198 13 839.3
PI THOMAS SPELT, BERND HINTZMAN, SCHRITT ARMIN, CHRISTIAN PIRARSKI,
PI DUHL EDGAR,
PI ANDRE ROSENTHAL
PC C12N15/09, A61K31/711, A61K38/00, A61K48/00, A61P35/00, C07K14/47,
PC C07K16/18,
PC C12N1/19, C12N5/10, C12N5/00, A61K37/02, C12N5/00 CC Human
nucleic acid sequence originating in mammary tumor CC
tissue
FT Key Location/Qualifiers
FH source l. .2281
FT /organism="Homo sapiens (human)".

```

```

FEATURES
source
1. .2281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

BASE COUNT 601 a 498 c 494 g 688 t
ORIGIN

```

```

Query Match 99.7%; Score 499.4; DB 6; Length 2281;
Best Local Similarity 99.8%; Pred. No. fe-136;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCGCAAGACACCTCTCTGGAGGGCTCTCTGC 60
Db 880 CTCTCCAGCTGAACCAAAATATTACAAATTCGCGCAAGACACCTCTCTGGAGGGCTCTCTGC 939
QY 61 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 120
Db 940 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 999
QY 121 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
Db 1000 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1059
QY 181 ACCCAATGTAGTGGGGCCCAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 240
Db 1060 ACCCAATGTAGTGGGGCCCAAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1119
QY 241 CTGTGCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 300
Db 1120 CTGTGCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 1179
QY 301 AAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCGCTCCCAAGGATGAGGATGAGGATGAGG 360
Db 1180 AAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCGCTCCCAAGGATGAGGATGAGGATGAGG 1239
QY 361 CCCTGTTAGGCTTGGTTTAAAGATTTTATGTTGTTTTCGAATGTTTATGTTAGGCT 420
Db 1240 CCCTGTTAGGCTTGGTTTAAAGATTTTATGTTGTTTTCGAATGTTTATGTTAGGCT 1299
QY 421 TTTGGTGGAGCGGCGGAGCTGACAGGATCTTACAAGAGAAATTTGACATCTCTGGAAGC 480
Db 1300 TTTGGTGGAGCGGCGGAGCTGACAGGATCTTACAAGAGAAATTTGACATCTCTGGAAGC 1359
QY 481 TTAGCAATCTTATTGCACACT 501
Db 1360 TTAGCAATCTTATTGCACACT 1380

```

```

RESULT 9
ARI51390
LOCUS Sequence 6 from patent US 6232077.
DEFINITION
ACCESSION ARI51390
VERSION ARI51390.1 GI:15117440
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2311)
AUTHORS Au-Young, J., Guegler, K.J. and Hawkins, P.R.
TITLE Human protein kinases
JOURNAL Patent: US 6232077-A 6 15-MAY-2001;
FEATURES
Source
1. .2311
/organism="unknown"
BASE COUNT 604 a 508 c 506 g 692 t 1 others
ORIGIN

```

```

Query Match 99.7%; Score 499.4; DB 6; Length 2311;
Best Local Similarity 99.8%; Pred. No. fe-136;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCGCAAGACACCTCTCTGGAGGGCTCTCTGC 60
Db 947 CTCTCCAGCTGAACCAAAATATTACAAATTCGCGCAAGACACCTCTCTGGAGGGCTCTCTGC 1006
QY 61 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 120
Db 1007 AGAAGGACAGGCAAAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAAGATCATG 1086

```

QY 121 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 180
 DB 1067 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 1126
 QY 181 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 240
 DB 1127 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 1186
 QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCTCAAGG 300
 DB 1187 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCTCAAGG 1246
 QY 301 AAGTGGCAGGCTTTCTTAGGCTTTTCTATGCGCTCCGAGGACTCTTCTCTGAA 360
 DB 1247 AAGTGGCAGGCTTTCTTAGGCTTTTCTATGCGCTCCGAGGACTCTTCTCTGAA 1306
 QY 361 CCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTGTAGTACCT 420
 DB 1307 CCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTGTAGTACCT 1366
 QY 421 TTGTGGAGCCGACGACGACATCTTACAGAGATTTGACAGCTCTGAGAGC 480
 DB 1367 TTGTGGAGCCGACGACGACATCTTACAGAGATTTGACAGCTCTGAGAGC 1426
 QY 481 TTAGCAATCTATTGCACACT 501
 DB 1427 TTAGCAATCTATTGCACACT 1447

RESULT 10
 AR270209
 LOCUS Sequence 772 from patent US 6500938.
 DEFINITION AR270209
 ACCESSION AR270209
 VERSION AR270209.1 GI:29701443
 KEYWORDS
 ORGANISM Unknown.
 SOURCE Unclassified.
 REFERENCE 1 (bases 1 to 2311)
 AUTHORS Au-Young, J. and Seilhamer, J. J.
 TITLE Composition for the detection of signaling pathway gene expression
 JOURNAL Patent: US 6500938-A 772 31-DEC-2002;
 FEATURES
 source Location/Qualifiers
 1..2311
 BASE COUNT 604 a 508 c 506 g 692 t 1 others
 ORIGIN

Query Match 99.7%; Score 499.4; DB 6; Length 2311;
 Best Local Similarity 99.8%; Pred. No. 1e-136;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 60
 DB 947 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 1006
 QY 61 AGAGGACAGCACAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGATCATG 120
 DB 1007 AGAGGACAGCACAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGATCATG 1066
 QY 121 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 180
 DB 1067 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 1126
 QY 181 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 240
 DB 1127 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 1186
 QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCTCAAGG 300
 DB 1187 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCTCAAGG 1246

QY 301 AAGTGGCAGGCTTTCTTAGGCTTTTCTATGCGCTCCGAGGACTCTTCTCTGAA 360
 DB 1247 AAGTGGCAGGCTTTCTTAGGCTTTTCTATGCGCTCCGAGGACTCTTCTCTGAA 1306
 QY 361 CCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTGTAGTACCT 420
 DB 1307 CCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTGTAGTACCT 1366
 QY 421 TTGTGGAGCCGACGACGACATCTTACAGAGATTTGACAGCTCTGAGAGC 480
 DB 1367 TTGTGGAGCCGACGACGACATCTTACAGAGATTTGACAGCTCTGAGAGC 1426
 QY 481 TTAGCAATCTATTGCACACT 501
 DB 1427 TTAGCAATCTATTGCACACT 1447

RESULT 11
 BD021928
 LOCUS Human protein kinases.
 DEFINITION BD021928
 ACCESSION BD021928
 VERSION BD021928.1 GI:22563151
 KEYWORDS JP 2001506843-A/3.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 2311)
 AUTHORS Young, J. O., Segura, K. J. and Horkins, P. R.
 TITLE Human protein kinases
 JOURNAL Patent: JP 2001506843-A 3 29-MAY-2001;
 COMMENT INSITE PHARMACEUTICALS INC
 PN JP 2001506843-A/3
 PD 29-MAY-2001
 PF 10-SEP-1997 JP 1998513776
 PR 12-SEP-1996 US 08/712709
 PI JANICE O YOUNG, KARL J SEGURA, PHILLIP R HORKINS PC
 C12N15/00, A61K39/395, A61K39/395, A61K39/395, A61K48/00, A61K48/00, PC
 A61P11/06,
 PC A61P19/02, A61P25/28, A61P29/00, A61P35/00, A61P43/00, C07K16/40,
 C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/12, C12Q1/48, C12N15/00, C12N5/
 PC 00, A61K37/52
 CC
 FH

FEATURES
 source Location/Qualifiers
 1..2311
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 BASE COUNT 604 a 507 c 507 g 692 t 1 others
 ORIGIN

Query Match 99.7%; Score 499.4; DB 6; Length 2311;
 Best Local Similarity 99.8%; Pred. No. 1e-136;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 60
 DB 947 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 1006
 QY 61 AGAGGACAGCACAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGATCATG 120
 DB 1007 AGAGGACAGCACAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGATCATG 1066
 QY 121 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 180
 DB 1067 TCTTCTCTCTTAATTAAGTGGATGATCTCATTAATAAGAGATTAATCTCCCTTTA 1126
 QY 181 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 240
 DB 1127 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 1186
 QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCTCAAGG 300
 DB 1127 ACCCAATGTAGTGGCCCAAGCAGCTAGGACCTTTGACCCCGAGTTTACCGAGAGC 1186

```
QY 241 CTGCCCCAATCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
|||||
DB 1187 CTGCCCCAATCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 1246
|||||
QY 301 AAGTCGCGAGGCTTTCTAGGCTTTTCTATGCGCCCTCCACAGGACTCTTTCCCTGAA 360
|||||
DB 1247 AAGTCGCGAGGCTTTCTAGGCTTTTCTATGCGCCCTCCACAGGACTCTTTCCCTGAA 1306
|||||
QY 361 CCTCTTAGGCTGCTTTTAAAGGATTTTATGCTGTTTCCGAATGTTTAGTTAGGCT 420
|||||
DB 1307 CCTCTTAGGCTGCTTTTAAAGGATTTTATGCTGTTTCCGAATGTTTAGTTAGGCT 1366
|||||
QY 421 TTGCTGAGCGCGCCAGCTGACAGGACATCTTACAAGAGATTTTCACATCTCTCGAAGC 480
|||||
DB 1367 TTGCTGAGCGCGCCAGCTGACAGGACATCTTACAAGAGATTTTCACATCTCTCGAAGC 1426
|||||
QY 481 TTAGCAATCTTATTGCACACT 501
|||||
DB 1427 TTAGCAATCTTATTGCACACT 1447
|||||

RESULT 12
AX553549
LOCUS
DEFINITION Sequence 1 from Patent WO2074987.
ACCESSION AX553549
VERSION AX553549.1 GI:25897549
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Lang,F., Busjahn,A. and Luft,F.C.
TITLE Quantitative diagnostic analysis of hypertension
JOURNAL Patent:WO 02074987-A 1 26-SEP-2002;
Lang, Florian (DB)
FEATURES
source
location/Qualifiers
1..2354
/organism="Homo sapiens"
/db_type="genomic DNA"
/db_xref="taxon:9606"
43..1338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD58123.1"
/db_xref="GI:25897550"
/translation="MTVKTEAAKGLTYSRMGMVAILLAFMKORRMGLNDFIOKIAN
NSYACKHPEVQSILKISQOPELMNANPPSPSQINLGPSSNHPKPSDFHLK
VIGSGFGKVLARHKEVEFYAVKVLQKAILKKKEKHIMSERNVLLKRVKHPFLV
GLHFSFQADKLYFDVYINGGELFYHLQRECFLEPRARFYAAEIASALGYLHSLNI
VYDLKPENILLOSGHVLTFDGLCKENIEHNSTFTFCGTPYLAPEVLHKKOPYDR
TVDWMLGAVLYEMLYGLPPYSRNTAEYDNLNKLPLQKPNITLNSARHLEGLLOK
DRTKRLGAKDDFMEIKSHVFFSLINWDDLINKITPPENPNVSGNDLRHDPFTEF
PVPNSIGKSPDVLVTSVKEAEAFGLGFSVAPDSEF"
762
/note="1. SNP (C in T), stumme Mutation, d.h. beide
Versionen des SNPs resultieren in der AMINO ACID Asp in
der AMINO ACID- Position 240"
BASE COUNT 613 a 518 c 513 g 710 t
ORIGIN

Query Match 99.7%; Score 499.4; DB 6; Length 2354;
Best Local Similarity 99.8%; Pred. No. 1e-136;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATTACAATTCGCAAGACACTCTCTGAGGCGCTCTCTCC 60
|||||
DB 980 CTCTCCAGCTGAACCAATATTACAATTCGCAAGACACTCTCTGAGGCGCTCTCTCC 1039
|||||
QY 61 AGAAGGACAGCAAGCGGCTCGGGCCAAAGGATGACTTTCATGAGATTAAAGTCTATG 120
|||||

Db 1040 AGAAGGACAGCAAGCGGCTCGGGCCAAAGGATGACTTTCATGAGATTAAAGTCTATG 1099
QY 121 TCTTCTTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
|||||
DB 1100 TCTTCTTCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159
|||||
QY 181 ACCCAATATGTAGTGGGCGCCCAAGGAGCTAGGCACTTTGACCCCGAGTTTACCGAAGAGC 240
|||||
DB 1160 ACCCAATATGTAGTGGGCGCCCAAGGAGCTAGGCACTTTGACCCCGAGTTTACCGAAGAGC 1219
|||||
QY 241 CTGTCGCCAATCTTCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
|||||
DB 1220 CTGTCGCCAATCTTCCATTCGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 1279
|||||
QY 301 AAGTCGCGAGGCTTTCTAGGCTTTTCTATGCGCCCTCCACAGGACTCTTTCCCTGAA 360
|||||
DB 1280 AAGTCGCGAGGCTTTCTAGGCTTTTCTATGCGCCCTCCACAGGACTCTTTCCCTGAA 1339
|||||
QY 361 CCTCTTAGGCTGCTTTTAAAGGATTTTATGCTGTTTCCGAATGTTTAGTTAGGCT 420
|||||
DB 1340 CCTCTTAGGCTGCTTTTAAAGGATTTTATGCTGTTTCCGAATGTTTAGTTAGGCT 1399
|||||
QY 421 TTGCTGAGCGCGCCAGCTGACAGGACATCTTACAAGAGATTTTCACATCTCTCGAAGC 480
|||||
DB 1400 TTGCTGAGCGCGCCAGCTGACAGGACATCTTACAAGAGATTTTCACATCTCTCGAAGC 1459
|||||
QY 481 TTAGCAATCTTATTGCACACT 501
|||||
DB 1460 TTAGCAATCTTATTGCACACT 1480
|||||

RESULT 13
AK098509
LOCUS
DEFINITION Homo sapiens CDNA FLJ25643 fis, clone STM07148, highly similar to
SERINE/THREONINE-PROTEIN KINASE SGK (EC 2.7.1.-).
ACCESSION AK098509
VERSION AK098509.1 GI:21758535
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S.,
Morinaga,M., Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Nishikawa,T., Sugiyama,A., Kawakami,B., Negai,K., Isogai,T. and
Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2364)
AUTHORS Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing; Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing; RAB; clone selection for
full insert sequencing; RAB and Helix Research Institute.
FEATURES
source
location/Qualifiers
1..2364
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
```

```

/clone="STM07148"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/clone="cloning vector: pME18SFL3"
/notes="cloning vector: pME18SFL3"
BASE COUNT      615 a   524 c   517 g   708 t
ORIGIN

Query Match      99.7%; Score 499.4; DB 9; Length 2364;
Best Local Similarity 99.8%; Pred. No. 1e-136;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATACAAATTCGCGAGACACCTCTCTGGAGGCGCTCTCTGC 60
DB 995 CTCTCCAGCTGAACCAATATACAAATTCGCGAGACACCTCTCTGGAGGCGCTCTCTGC 1054
QY 61 AGAAGGACGAGCAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGATCATG 120
DB 1055 AGAAGGACGAGCAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGATCATG 1114
QY 121 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 1115 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1174
QY 181 ACCCAATGTAGTGGGCGCCCAAGCAGCTAGGCGACTTTGACCCGAGTTTACCGAGAGC 240
DB 1175 ACCCAATGTAGTGGGCGCCCAAGCAGCTAGGCGACTTTGACCCGAGTTTACCGAGAGC 1234
QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
DB 1235 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 1294
QY 301 AAGTCCGAGGCTTTCTTAGGCTTTTCTATGCGGCTCTCTGACGAGCTCTTCTCTGAA 360
DB 1295 AAGTCCGAGGCTTTCTTAGGCTTTTCTATGCGGCTCTCTGACGAGCTCTTCTCTGAA 1354
QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAGTGTCTTAGTAGGCT 420
DB 1355 CCCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAGTGTCTTAGTAGGCT 1414
QY 421 TTTGGTGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 480
DB 1415 TTTGGTGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 1474
QY 481 TTAGCAATCTTATTCACACT 501
DB 1475 TTAGCAATCTTATTCACACT 1495

RESULT 14
AF153609 Homo sapiens serine/threonine protein kinase sgk mRNA, complete cds.
LOCUS AF153609 2382 bp mRNA linear PRI 28-JUN-1999
DEFINITION Homo sapiens serine/threonine protein kinase sgk mRNA, complete cds.
ACCESSION AF153609
VERSION AF153609.1 GI:5231142
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Kim, M.K., Kim, Y.H., Seo, J.M., Lee, H.M., Chung, H.J., Sohn, M.Y., Hwang, S.Y., Im, S.O., Jung, E.J. and Kim, J.C.
A catalogue of genes in the human dermal papilla cells as identified by expressed sequence tags
Unpublished
2 (bases 1 to 2382)
Kim, M.K., Kim, Y.H., Suh, J.M., Lee, H.M., Chung, H.J., Sohn, M.Y., Hwang, S.Y., Im, S.O., Jung, E.J. and Kim, J.C.
Direct Submission
Submitted (24-MAY-1999) Immunology, Kyungpook National University, School of Medicine, 101 Dongin Dong, Jung Gu, Taegu, Taegu 700-422, South Korea

/clone="STM07148"
/tissue_type="stomach mucosa"
/clone_lib="STM"
/clone="cloning vector: pME18SFL3"
/notes="cloning vector: pME18SFL3"
BASE COUNT      615 a   524 c   517 g   708 t
ORIGIN

Query Match      99.7%; Score 499.4; DB 9; Length 2364;
Best Local Similarity 99.8%; Pred. No. 1e-136;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATACAAATTCGCGAGACACCTCTCTGGAGGCGCTCTCTGC 60
DB 995 CTCTCCAGCTGAACCAATATACAAATTCGCGAGACACCTCTCTGGAGGCGCTCTCTGC 1054
QY 61 AGAAGGACGAGCAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGATCATG 120
DB 1055 AGAAGGACGAGCAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGATCATG 1114
QY 121 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 1115 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1174
QY 181 ACCCAATGTAGTGGGCGCCCAAGCAGCTAGGCGACTTTGACCCGAGTTTACCGAGAGC 240
DB 1175 ACCCAATGTAGTGGGCGCCCAAGCAGCTAGGCGACTTTGACCCGAGTTTACCGAGAGC 1234
QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
DB 1235 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 1294
QY 301 AAGTCCGAGGCTTTCTTAGGCTTTTCTATGCGGCTCTCTGACGAGCTCTTCTCTGAA 360
DB 1295 AAGTCCGAGGCTTTCTTAGGCTTTTCTATGCGGCTCTCTGACGAGCTCTTCTCTGAA 1354
QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAGTGTCTTAGTAGGCT 420
DB 1355 CCCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAGTGTCTTAGTAGGCT 1414
QY 421 TTTGGTGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 480
DB 1415 TTTGGTGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 1474
QY 481 TTAGCAATCTTATTCACACT 501
DB 1475 TTAGCAATCTTATTCACACT 1495

RESULT 15
BC001263
LOCUS BC001263 2382 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, serum/glucocorticoid regulated kinase, clone MGC:5030 IMAGE:3459056, mRNA, complete cds.
ACCESSION BC001263
VERSION BC001263.1 GI:12654838
KEYWORDS MGC.
SOURCE Homo sapiens (human)

```

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2382)
 Strausberg, R.
 Direct Submission
 Submitted (11-DEC-2000) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickson@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: TRAX Plate: 4 Row: a Column: 12
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5032090.

FEATURES
 source
 1..2382
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="LocusID:6446"
 /db_xref="taxon:9606"
 /clone="MGC:5030 IMAGE:3459056"
 /tissue_type="Cervix, carcinoma"
 /clone_lib="NTH_MGC_12"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 43..1338
 /codon_start=1
 /product="serum/glucocorticoid regulated kinase"
 /protein_id="AAH01283.1"
 /db_xref="GI:12654839"

CDS

/translation="MTVTEAAKGTITYSRMGNVAILAFMKQRRMGLNDFQKIAN
 NSYACKPEVOSILKISQPELMANPPSPSQINLGPSSNPFAKPSDFHFLK
 VIGKSGKVLARHKADEVYAVKLOKKAALKKEERHIMSERVLLKNVKKHFLV
 GLHFSOTADLYPVDYINGGELFYHLORECFLEPRAFYAAETASALGYLHSLNI
 VYDLKPEILLDSQGHVLTDFGLCKENIEHNSSTTCGTPPEYLAPVLEHQPYDR
 TVDWCLGNLYLGLPFFYSERNAENDNINLPLQLKPNITNSARHLEGLLQK
 DTRRLGAKDDFMEIKSHVFSLINMDLLINKKITPPFPNVSGPNDLRHFDPEFTEE
 PVPNSIGKSPDSVLVTASVKEAAPAFIGFSYAPPTDSFL"

BASE COUNT 642 a 519 c 513 g 708 t
 ORIGIN

Query Match 99.7%, Score 499.4; DB 9; Length 2382;
 Best Local Similarity 99.8%, Pred. No. 1e-136;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTCCTCCAGCTGAACCAATATTACAAATTCGGAAGACACCTCCTGGAGGCGCTCTGC 60
 DB 980 CTCCTCCAGCTGAACCAATATTACAAATTCGGAAGACACCTCCTGGAGGCGCTCTGC 1039
 QY 61 AGAAGGAGGACCAAGCGGCTCGGGCCAGGATGACTTCATGGAGATTAAAGATCATG 120
 DB 1040 AGAAGGAGGACCAAGCGGCTCGGGCCAGGATGACTTCATGGAGATTAAAGATCATG 1099
 QY 121 TCTTCTCTCTCTTAATTAACCTGGATGATCTCATTATAAGAGATTACTCCCTTTTA 180
 DB 1100 TCTTCTCTCTCTTAATTAACCTGGATGATCTCATTATAAGAGATTACTCCCTTTTA 1159
 QY 181 ACCCAATGTGAGTGGGCGGCAACAGAGCTACGGCACTTTGACCGGAGTTTACCGAAGAGC 240

Db 1160 ACCCAATGTGAGTGGGCGGCAACAGAGCTACGGCACITTTGACCCCGAGTTTACCGAAGAGC 1219
 QY 241 CTGTCCCGCAKCTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
 DB 1220 CTGTCCCGCAKCTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 1279
 QY 301 AAGCTGGCGAGGCTTTCTAGGCTTTTCTATGGCGCTCCCGACGGAGCTCTTTCTTGA 360
 DB 1280 AAGCTGGCGAGGCTTTCTAGGCTTTTCTATGGCGCTCCCGACGGAGCTCTTTCTTGA 1339
 QY 361 CCCTGTAGGCGTGGTGTAAAGGATTTATCTGTGTGTTCCGAATGTTTTAGTTAGCT 420
 DB 1340 CCCTGTAGGCGTGGTGTAAAGGATTTATCTGTGTGTTCCGAATGTTTTAGTTAGCT 1399
 QY 421 TTTGTGTAGGCGGCGGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 480
 DB 1400 TTTGTGTAGGCGGCGGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGC 1459
 QY 481 TTAGCAATCTTTATTGGCACT 501
 DB 1460 TTAGCAATCTTTATTGGCACT 1480

Search completed: August 11, 2003, 14:24:55
 Job time : 2018 secs

XX WPI: 1998-449109/39.
 XX P-PSDB; AAM77219.
 XX New nucleic acid encoding cell-volume regulating kinase h-sgk and
 XX related proteins - used for diagnosis and treatment of diseases
 XX involving changes in cell volume, e.g. renal insufficiency,
 XX inflammation, infections etc.
 XX Disclosure: Fig 1; 15pp; German.
 XX The human cell-volume regulating kinase h-sgk is inhibited by the
 XX swelling of cells (or presence of urea), whereas cell shrinkage
 XX stimulates its expression. The nucleic acid h-sgk, and fragments, are
 XX particularly used to detect changes in cell volume, specifically for
 XX diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
 XX natraemia, diabetes mellitus, renal insufficiency, hypercatabolism,
 XX hepatic encephalopathy, inflammation, microbial/viral infection, fructose
 XX intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
 XX The nucleic acid, protein and products including receptors that bind
 XX h-sgk, can be used to treat these disorders.
 XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
 XX
 XX Query Match 100.0%; Score 501; DB 19; Length 2370;
 XX Best Local Similarity 100.0%; Pred. No. 5.2e-153;
 XX Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 CTCCTCAGCTGAACCAATATATACAAATTCGCAAGACACCTCTCGAGGGCTCTCTGC 60
 XX 980 CTCCTCAGCTGAACCAATATATACAAATTCGCAAGACACCTCTCGAGGGCTCTCTGC 1039
 XX
 XX 61 AGAAGACAGGACAAAGCGCTCGGGCCCAAGAGTATCTTCATGAGATTAAAGATCATG 120
 XX 1040 AGAAGACAGGACAAAGCGCTCGGGCCCAAGAGTATCTTCATGAGATTAAAGATCATG 1099
 XX
 XX 121 TCCTCTCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
 XX 1100 TCCTCTCTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159
 XX
 XX 181 ACCCAATATGAGTGGGCGCCACAGCTACGACCTTTGACCCCGAGTTTACCGAAGAGC 240
 XX 1160 ACCCAATATGAGTGGGCGCCACAGCTACGACCTTTGACCCCGAGTTTACCGAAGAGC 1219
 XX
 XX 241 CTGTCCTCCCACTTCATGCGAATGCGCTGACAGCTCTCTGACAGCGGTCAAGC 300
 XX 1220 CTGTCCTCCCACTTCATGCGAATGCGCTGACAGCTCTCTGACAGCGGTCAAGC 1279
 XX
 XX 301 AAGCTCCGAGGCTTCTTCTAGGCTTTTCTATGCGCCCTCCACGAGCTCTTCTCTGAA 360
 XX 1280 AAGCTCCGAGGCTTCTTCTAGGCTTTTCTATGCGCCCTCCACGAGCTCTTCTCTGAA 1339
 XX
 XX 361 CCTGTAGGCTTGTGTTTAAAGGNTTTATGTTGTTTCCGAATGTTTGTAGTACCT 420
 XX 1340 CCTGTAGGCTTGTGTTTAAAGGNTTTATGTTGTTTCCGAATGTTTGTAGTACCT 1399
 XX
 XX 421 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAAGAAATTTGACATCTCTGGAAGC 480
 XX 1400 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAAGAAATTTGACATCTCTGGAAGC 1459
 XX
 XX 481 TTAGCAATCTTATTGACACT 501
 XX 1460 TTAGCAATCTTATTGACACT 1480
 XX
 XX RESULT 2
 XX ID ABK84749
 XX AC ABK84749 standard; cDNA; 2370 BP.
 XX DT 14-AUG-2002 (first entry)
 XX

DE Human cDNA differentially expressed in granulocytic cells #120.
 XX
 XX Human; ss: granulocytic cell; DNA chip; bacterial infection;
 XX viral infection; parasitic infection; protozoal infection;
 XX fungal infection; sterile inflammatory disease; psoriasis;
 XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
 XX adult respiratory distress syndrome; inflammatory bowel disease;
 XX Crohn's disease; ulcerative colitis; periodontal disease;
 XX granulocyte activation; chronic inflammation; allergy.
 XX
 XX Homo sapiens.
 XX WO200228999-A2.
 XX 11-APR-2002.
 XX 03-OCT-2001; 2001WO-US30821.
 XX 03-OCT-2000; 2000US-237189P.
 XX (GENE-) GENE LOGIC INC.
 XX Beazer-Barclay Y, Weissman SM, Yamaya S, Vockley J;
 XX WPI; 2002-435328/46.
 XX
 XX Detecting granulocyte activation by detecting differential expression
 XX of genes associated with granulocyte activation, which serves as
 XX diagnostic markers that is useful for monitoring disease states and
 XX drug toxicity
 XX
 XX Claim 1: SEQ ID No 1320; 114pp; English.
 XX
 XX The invention relates to detecting (M1) granulocyte (GC) activation
 XX (GCA) by detecting the level of expression of gene(s) (Gs) identified by
 XX DNA chip analysis as given in the specification, and comparing
 XX the expression level to an expression level in an unactivated
 XX GC, where differential expression of Gs is indicative of GCA.
 XX Also included are modulating (M2) GA by contacting GC with an agent
 XX that alters the expression of at least one gene in Gs; (2) screening (M3)
 XX for an agent capable of modulating GCA or an inflammation (especially
 XX chronic) in a tissue, an allergic response in a subject, exposure of a
 XX subject to a pathogen or sterile inflammatory disease using the
 XX gene expression profile; (3) detecting (M4) an inflammation (especially
 XX chronic) in a tissue, an allergic response in a subject, exposure of a
 XX subject to a pathogen or sterile inflammatory disease, by detecting the
 XX level of expression in a sample of the tissue of gene(s) from Gs, where
 XX the level of expression of the gene is indicative of inflammation;
 XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,
 XX an allergic response in a subject, exposure of a subject to a pathogen
 XX or sterile inflammatory disease, by contacting a tissue having
 XX inflammation with an agent that modulates the expression of gene(s)
 XX from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
 XX modulating GA; M3 is useful for screening an agent capable of modulating
 XX GCA preferably in an inflammation in a tissue; M4 is useful for
 XX detecting an inflammation (especially chronic) in a tissue, or allergic
 XX response in a subject, exposure of a subject to a pathogen or sterile
 XX inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
 XX glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
 XX reperfusion injury, ARDS, adult respiratory distress syndrome,
 XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 XX parasitic infection, protozoal infection, viral infection,
 XX periodontal disease; also bacterial infection, fungal infection, and M5 is
 XX useful for treating one of the above conditions. The present
 XX sequence represents a gene differentially expressed in granulocytes.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
 XX

Query Match 100.0%; Score 501; DB 24; Length 2370;
 Best Local Similarity 100.0%; Pred. No. 5.2e-153;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 60
 DB 980 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 1039

QY 61 AGAAGGACAGGACAAACCGCTCGGGCCCAAGAGATGCTTCATGGAGATTAAGAGTCATG 120
 DB 1040 AGAAGGACAGGACAAACCGCTCGGGCCCAAGAGATGCTTCATGGAGATTAAGAGTCATG 1099

QY 121 TCTTCT 180
 DB 1100 TCTTCT 1159

QY 181 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 240
 DB 1160 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 1219

QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 300
 DB 1220 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 1279

QY 301 AGCTGCCAGGCTTTCTCTAGGCTTTCTCTATGCGCTCCACGAGTTCACCAAGAGC 240
 DB 1160 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 1219

QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 300
 DB 1220 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 1279

QY 301 AGCTGCCAGGCTTTCTCTAGGCTTTCTCTATGCGCTCCACGAGTTCACCAAGAGC 240
 DB 1160 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 1219

QY 361 CCTGTCTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCCAATGTTTGTAGTACGCT 420
 DB 1340 CCTGTCTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCCAATGTTTGTAGTACGCT 1399

QY 421 TTGTGTGGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 480
 DB 1400 TTGTGTGGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 1459

QY 481 TTAGCAATCTTATTCACACT 501
 DB 1460 TTAGCAATCTTATTCACACT 1480

RESULT 3
 ABN97360
 ID ABN97360 standard; DNA; 2370 BP.
 AC ABN97360;
 DT 13-AUG-2002 (first entry)
 DE Gene #3858 used to diagnose liver cancer.

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 metastatic liver tumor; cytostatic; expression profile; disease state;
 disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.
 XX W0200229103-A2.
 PN 11-APR-2002.
 XX 02-OCT-2001; 2001WO-US30589.
 XX 02-OCT-2000; 2000US-237054P.
 PR (GENE-) GENE LOGIC INC.
 XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
 PI WPI; 2002-426119/45.
 XX Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a
 liver tissue sample -
 PS Claim 1; SEQ ID NO 3858; 298pp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN93503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatotropic, and
 CC cytostatic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from Wipo
 CC at ftp.wipo.int/pub/published_pat_sequences.

Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Query Match 100.0%; Score 501; DB 24; Length 2370;
 Best Local Similarity 100.0%; Pred. No. 5.2e-153;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 60
 DB 980 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCTCTCTGC 1039

QY 61 AGAAGGACAGGACAAACCGCTCGGGCCCAAGAGATGCTTCATGGAGATTAAGAGTCATG 120
 DB 1040 AGAAGGACAGGACAAACCGCTCGGGCCCAAGAGATGCTTCATGGAGATTAAGAGTCATG 1099

QY 121 TCTTCT 180
 DB 1100 TCTTCT 1159

QY 181 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 240
 DB 1160 ACCCAATGTAGTGGGCCCAAGAGCTAGGACCTTTGACCCGAGTTCACCAAGAGC 1219

QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 300
 DB 1220 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGTCCCTGACAGCGTCCCTGACAGCGTCAAGG 1279

QY 301 AGCTGCCAGGCTTTCTCTAGGCTTTCTCTATGCGCTCCACGAGTTCACCAAGAGC 240
 DB 1280 AGCTGCCAGGCTTTCTCTAGGCTTTCTCTATGCGCTCCACGAGTTCACCAAGAGC 1339

QY 361 CCTGTCTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCCAATGTTTGTAGTACGCT 420
 DB 1340 CCTGTCTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCCAATGTTTGTAGTACGCT 1399

QY 421 TTGTGTGGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 480
 DB 1400 TTGTGTGGAGCCGCGAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 1459

QY 481 TTAGCAATCTTATTCACACT 501
 DB 1460 TTAGCAATCTTATTCACACT 1480

RESULT 4
 ABL70006
 ID ABL70006 standard; DNA; 2370 BP.
 XX ABL70006;
 AC ABL70006;
 XX 15-MAY-2002 (first entry)
 DT Pancreas cancer related gene sequence SEQ ID NO:8343.
 XX

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX Homo sapiens.

XX W0200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001W0-US10838.

XX 05-JUN-2000; 2000US-209473P.

XX 05-JUN-2000; 2000US-209531P.

XX 18-SEP-2000; 2000US-233133P.

XX 18-SEP-2000; 2000US-233617P.

XX 20-SEP-2000; 2000US-234009P.

XX 20-SEP-2000; 2000US-234034P.

XX 20-SEP-2000; 2000US-234052P.

XX 22-SEP-2000; 2000US-234509P.

XX 22-SEP-2000; 2000US-234567P.

XX 25-SEP-2000; 2000US-234923P.

XX 25-SEP-2000; 2000US-234924P.

XX 25-SEP-2000; 2000US-235077P.

XX 25-SEP-2000; 2000US-235082P.

XX 25-SEP-2000; 2000US-235113P.

XX 25-SEP-2000; 2000US-235280P.

XX 26-SEP-2000; 2000US-235637P.

XX 26-SEP-2000; 2000US-235638P.

XX 27-SEP-2000; 2000US-235711P.

XX 27-SEP-2000; 2000US-235720P.

XX 27-SEP-2000; 2000US-235840P.

XX 27-SEP-2000; 2000US-235863P.

XX 28-SEP-2000; 2000US-236028P.

XX 28-SEP-2000; 2000US-236032P.

XX 28-SEP-2000; 2000US-236033P.

XX 28-SEP-2000; 2000US-236034P.

XX 28-SEP-2000; 2000US-236109P.

XX 28-SEP-2000; 2000US-236111P.

XX 29-SEP-2000; 2000US-236842P.

XX 29-SEP-2000; 2000US-236891P.

XX 02-OCT-2000; 2000US-237172P.

XX 02-OCT-2000; 2000US-237173P.

XX 02-OCT-2000; 2000US-237278P.

XX 02-OCT-2000; 2000US-237294P.

XX 02-OCT-2000; 2000US-237295P.

XX 02-OCT-2000; 2000US-237318P.

XX 03-OCT-2000; 2000US-237425P.

XX 03-OCT-2000; 2000US-237598P.

XX 03-OCT-2000; 2000US-237604P.

XX 03-OCT-2000; 2000US-237606P.

XX 03-OCT-2000; 2000US-237608P.

XX 01-NOV-2000; 2000US-244867P.

XX 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

XX Soppet DR, Weaver Z;

XX WPI: 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -

XX Claim 1; SEQ ID 8343; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in

CC expression of at least one gene (1) of a signature gene set, where (1)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL51064
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;

Query Match 100.0%; Score 501; DB 24; length 2370;

Best Local Similarity 100.0%; bred. No. 5, 2e-153;

Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATATACAAATTCGCAAGACACCTCGTGGAGCGCTCTGCG 60
 DB 980 CTCTCCAGCTGAACCAATATATACAAATTCGCAAGACACCTCGTGGAGCGCTCTGCG 1039
 QY 61 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAGAGCATG 120
 DB 1040 AGAAGACAGGACAAAGCGCTCGGGCCCAAGATGACTTCATGGAGATTAAGAGCATG 1099
 QY 121 TCTCTCTCTCTTAATTAAGTGGGATGCTCTATTAAGAGATTAAGTCTCTCTCTTA 180
 DB 1100 TCTCTCTCTCTTAATTAAGTGGGATGCTCTATTAAGAGATTAAGTCTCTCTCTTA 1159
 QY 181 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGCACTTTGACCCGAGTTTACGAGAGG 240
 DB 1160 ACCCAATGTGAGTGGGCCCAAGAGCTAGCGCACTTTGACCCGAGTTTACGAGAGG 1219
 QY 241 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGCTCTGTCACAGCAGTGTCAAGG 300
 DB 1220 CTGTCCCAACTCCATTTGGCAAGTCCCTGACAGCGCTCTGTCACAGCAGTGTCAAGG 1279
 QY 301 AAGCTGCGGAGGCTTCTCTAGGCTTTTCCATCGGCTCCCGAGGACTCTTCTCTCTAA 360
 DB 1280 AAGCTGCGGAGGCTTCTCTAGGCTTTTCCATCGGCTCCCGAGGACTCTTCTCTCTAA 1339
 QY 361 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTATGTTAGTCT 420
 DB 1340 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTATGTTAGTCT 1399
 QY 421 TTGGTGGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGAGAGC 480
 DB 1400 TTGGTGGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGAGAGC 1459
 QY 481 TTAGCAATCTTATGACACT 501
 DB 1460 TTAGCAATCTTATGACACT 1480

RESULT 5

AAZ33647

ID AAZ33647 standard; cDNA: 2281 BP.

XX

AC AAZ33647;

XX

DT 08-DEC-1999 (first entry)

XX

DE Human breast tumour-associated EST 37.

XX

KW Expressed sequence tag; EST; human; breast; cancer; gene therapy;

KW treatment; tumour; cytostatic; medicament; ss.

XX

OS Homo sapiens.

XX

PN DEL9813839-AL.
 XX 23-SEP-1999.
 XX 20-MAR-1998; 98DE-1013839.
 XX 20-MAR-1998; 98DE-1013839.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E, Rosentahl A;
 PI WPI; 1999-528981/45.
 XX P-PSDB; AAY48573.
 XX Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -
 XX Claim 1a; 116-117; 188pp; German.
 XX This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AAZ33611-48617 represents expressed
 CC sequence tags described in the method of the invention.
 XX Sequence 2281 BP; 601 A; 498 C; 494 G; 688 T; 0 other;
 SQ

Query Match 99.7%; Score 499.4; DB 20; Length 2281;
 Best Local Similarity 99.8%; Pred. No. 1.7e-152;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGAGGGCTCTCTGC 60
 DB 880 CTCCTCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGAGGGCTCTCTGC 939
 QY 61 AGAAGCAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATG 120
 DB 940 AGAAGCAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATG 999
 QY 121 TCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTATAGAAGATTACTCCCTTTTA 180
 DB 1000 TCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTATAGAAGATTACTCCCTTTTA 1059
 QY 181 ACCCAATGTGAGTGGGCGCCCAAGCAGCTACGGCCTTTGACCCGAGTTTACCGAAGC 240
 DB 1060 ACCCAATGTGAGTGGGCGCCCAAGCAGCTACGGCCTTTGACCCGAGTTTACCGAAGC 1119
 QY 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCTCTCTGACAGCCAGGCTCAAGG 300
 DB 1120 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCTCTCTGACAGCCAGGCTCAAGG 1179
 QY 301 AAGCTGCGAGGCTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCT 360
 DB 1180 AAGCTGCGAGGCTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCT 1239
 QY 361 CCTCTGTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCT 420
 DB 1240 CCTCTGTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCTTTTCTTAGGCT 1299
 QY 421 TTTGGTGGAGCGCGCAGCTGACAGACATCTTACAAGAAATTTGCACATCTCTGGAGC 480
 DB 1300 TTTGGTGGAGCGCGCAGCTGACAGACATCTTACAAGAAATTTGCACATCTCTGGAGC 1359
 QY 481 TTAGCAATCTTATTCACACT 501
 DB 1360 TTAGCAATCTTATTCACACT 1380

RESULT 6

AAV23833
 ID AAV23833 standard; cDNA; 2311 BP.
 XX
 AC AAV23833;
 XX
 DT 31-JUL-1998 (first entry)
 XX
 DE Human protein kinase HPK-3 coding sequence.
 XX
 XX Protein kinase; human; HPK; signalling cascade; kinase expression;
 KW Alzheimer's disease; cancer; anti-inflammatory agent; immunosuppressor;
 KW asthma; multiple sclerosis; rheumatoid arthritis; lymphocytic leukaemia;
 KW lymphoma; therapy; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 379..1080
 FT /*taq- a
 XX
 XX WO9811234-A2.
 PN
 XX 19-MAR-1998.
 PD
 XX 10-SEP-1997; 97WO-US15923.
 PF
 XX 12-SEP-1996; 96US-0712709.
 PR
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Au-Young J, Guegler KJ, Hawkins PK;
 PI WPI; 1998-207394/18.
 DR P-PSDB; AAW54205.
 XX
 XX New isolated human protein kinase(s) - used to develop products for
 PT diagnosis and treatment of e.g. Alzheimer's disease, cancers,
 PT asthma, multiple sclerosis or rheumatoid arthritis
 XX
 XX Claim 25; Fig 3; 75pp; English.
 PS
 XX This sequence encodes a human protein kinase (HPK) of the invention. The
 CC HPK protein can be used to develop products for studying signalling
 CC cascades in various cells and tissues, diagnosing disease and selecting
 CC inhibitors or drugs with the potential to intervene in various disorders
 CC or diseases in which altered kinase expression is implicated. The
 CC products can be used to e.g. reverse memory loss such as due to
 CC Alzheimer's disease, for cancer treatment or to provide anti-inflammatory
 CC and immunosuppressive effects in the treatment of e.g. asthma, multiple
 CC sclerosis, rheumatoid arthritis, as well as certain cancers, e.g.
 CC lymphocytic leukaemias or lymphomas.
 XX
 SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;

Query Match 99.7%; Score 499.4; DB 19; Length 2311;
 Best Local Similarity 99.8%; Pred. No. 1.7e-152;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGAGGGCTCTCTGC 60
 DB 947 CTCCTCAGCTGAACCAATATTACAAATCCGCAAGACACCTCTCTGAGGGCTCTCTGC 1006
 QY 61 AGAAGCAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATG 120
 DB 1007 AGAAGCAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATG 1066
 QY 121 TCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTATAGAAGATTACTCCCTTTTA 180
 DB 1067 TCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTATAGAAGATTACTCCCTTTTA 1126
 QY 181 ACCCAATGTGAGTGGGCGCCCAAGCAGCTACGGCCTTTGACCCGAGTTTACCGAAGC 240
 DB 1127 ACCCAATGTGAGTGGGCGCCCAAGCAGCTACGGCCTTTGACCCGAGTTTACCGAAGC 1186

QY 241 CTGTCCCAACTCCATGTCAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAGG 300
 DB 1187 CTGTCCCAACTCCATGTCAGTCCCTGACAGCGTCTCTGACAGCCAGCGTCAAGG 1246
 QY 301 AAGCTGCCGAGGCTTTCCTAGGCTTTCCTATGCGCTCCACGAGACTCTTCCCTCTGAA 360
 DB 1247 AAGCTGCCGAGGCTTTCCTAGGCTTTCCTATGCGCTCCACGAGACTCTTCCCTCTGAA 1306
 QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATGTTAGCT 420
 DB 1307 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATGTTAGCT 1366
 QY 421 TTGGTGGAGCGCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 480
 DB 1367 TTGGTGGAGCGCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 1426
 QY 481 TTAGCAATCTTATTGCACACT 501
 DB 1427 TTAGCAATCTTATTGCACACT 1447

RESULT 7

ACA56174
 ID ACA56174 standard; cDNA; 2311 BP.

XX ACA56174;

DT 06-JUN-2003 (first entry)

XX Human signalling pathway polynucleotide probe SEQ ID NO 772.

DE Human; probe; ss; array element; Parkinson's disease;

KW signalling pathway population; cancer; adenocarcinoma; leukaemia;

KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

OS Homo sapiens.

XX US6500938-B1.

PN 31-DEC-2002.

XX 30-JAN-1998; 98US-0016434.

XX 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a
 PT microarray for monitoring the expression of a number of target
 PT polynucleotides

PS Claim 1; SEQ ID NO 772; 65pp; English.

XX The invention relates to a combination which, comprises a number of
 CC polynucleotide probes comprising a sequence selected from one of the 1490
 CC sequences mentioned in the specification. The combination is useful as an
 CC array element in a microarray for monitoring the expression of a number
 CC of target polynucleotides. The microarray is particularly useful in the
 CC diagnosis and treatment of cancer and immunopathology and neuropathology.
 CC The microarray is useful in diagnostics and treatment regimens, drug
 CC discovery and development, toxicological and carcinogenicity studies,
 CC forensics and pharmacogenomics. The microarray is also useful for
 CC monitoring progression of diseases and for developing sophisticated
 CC profiles for the effects of currently available therapeutic drugs. The
 CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
 CC and genomic fragments and in research and diagnostic applications. The
 CC array can detect changes in expression in a large number of genes coding
 CC for different signaling pathway populations which can be used to diagnose

CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
 CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
 CC and Parkinson's disease. The present sequence represents a polynucleotide
 CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?docId=06500938B1.

XX Sequence 2311 BP: 604 A; 508 C; 506 G; 692 T; 1 other;

Query Match 99.7%; Score 499.4; DB 25; Length 2311;
 Best Local Similarity 99.8%; Pred. No. 1.7e-152;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTGAACCAAAATATACAAATTCGCCAAGACACCTCTCTGGAGAGATCTCTG 60
 DB 947 CTCCTCAGCTGAACCAAAATATACAAATTCGCCAAGACACCTCTCTGGAGAGATCTCTG 1006
 QY 61 AGAAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTATG 120
 DB 1007 AGAAGGACAGGACAAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTATG 1066
 QY 121 TCTTCTCTCTTAATTAACCTGGATGATCTCATTAATAAGAGATTAACCTCTCTTA 180
 DB 1067 TCTTCTCTCTTAATTAACCTGGATGATCTCATTAATAAGAGATTAACCTCTCTTA 1126
 QY 181 ACCCAAAATGTGAGTGGCGCCCAAGCGCTAGGCGCACTTTGACCCCGAGTTTACGAAATG 240
 DB 1127 ACCCAAAATGTGAGTGGCGCCCAAGCGCTAGGCGCACTTTGACCCCGAGTTTACGAAATG 1186
 QY 241 CTGTCCCAACTCCATGTCAGTCCCTGACAGCGTCCCTGACAGCCAGCGTCAAGG 300
 DB 1187 CTGTCCCAACTCCATGTCAGTCCCTGACAGCGTCCCTGACAGCCAGCGTCAAGG 1246
 QY 301 AAGCTGCCGAGGCTTTCCTAGGCTTTCCTATGCGCTCCACGAGACTCTTCTCTGAA 360
 DB 1247 AAGCTGCCGAGGCTTTCCTAGGCTTTCCTATGCGCTCCACGAGACTCTTCTCTGAA 1306
 QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATGTTAGCT 420
 DB 1307 CCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCCGAATGTTTATGTTAGCT 1466
 QY 421 TTGGTGGAGCGCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 480
 DB 1367 TTGGTGGAGCGCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGC 1426
 QY 481 TTAGCAATCTTATTGCACACT 501
 DB 1427 TTAGCAATCTTATTGCACACT 1447

RESULT 8

ABX74395
 ID ABX74395 standard; cDNA; 2343 BP.

XX ABX74395;

XX 21-MAR-2003 (first entry)

DE Human cDNA sequence #6 up-regulated in non-aggressive CC-RCC.

XX Human; microarray; solid surface; immobilised probe; CC-RCC;
 KW differential expression profile; aggressive CC-RCC tumour type;
 KW non-aggressive CC-RCC tumour type; clear cell renal carcinoma;
 KW gene expression profiling; tumour tissue; gene; ss.

OS Homo sapiens.

XX WO200279411-A2.

XX 10-OCT-2002.

XX 29-MAR-2002; 2002WO-US09576.

XX PR 29-MAR-2001; 200105-279411P.
XX PA (VAND-) VAN ANDEL INST.
XX PI Haab B, Rhodes D, Teh BT, Takashi M;
XX PS WPI: 2003-040679/03.
XX
XX New microarray, comprising a matrix of cDNA probe from a set of probes
XX immobilised to a solid surface in predetermined order, useful in the
XX prognosis of patients with clear cell renal carcinoma -
XX
XX Claim 1: Page 79-80; 179pp; English.
XX
XX The present invention relates to a microarray comprising a matrix of
XX at least one cDNA probe from a set of probes immobilised to a solid
XX surface in a predetermined order, where a row of pixels corresponds
XX to replicates of one distinct probe from the set. The probes are
XX complementary to nucleic acid sequences that are expressed
XX differentially in aggressive as compared to non-aggressive types of
XX clear cell renal carcinoma (CC-RCC) and that hybridise to the probes
XX under high stringency conditions. The microarray is useful for the
XX prognosis of patients with CC-RCC, wherein aggressive and
XX non-aggressive CC-RCC tumour types are characterised by differential
XX expression profiles of genes that hybridise with one or more probes
XX immobilised on the microarray. The arrays are useful for gene
XX expression profiling of tumour and normal tissues. The present
XX sequence represents a human cDNA sequence up-regulated in
XX non-aggressive CC-RCC phenotypes.
XX
XX Sequence 2343 BP; 612 A; 511 G; 703 T; 0 other;

Query Match 99.7%; Score 499.4; DB 25; Length 2343;
Best Local Similarity 99.8%; Pred. No. 1.7e-152;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCTCAGCTGAACCAATATATACAAATTCGGCAAGACACCTCTCGAGGGCTCTCTGC 60
DB 973 CTCTCAGCTGAACCAATATATACAAATTCGGCAAGACACCTCTCGAGGGCTCTCTGC 1032
QY 61 AGAAGACAGGACCAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAGAGTCATG 120
DB 1033 AGAAGACAGGACCAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAGAGTCATG 1092
QY 121 TCTTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
DB 1093 TCTTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1152
QY 181 ACCCAATGTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGC 240
DB 1153 ACCCAATGTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGC 1212
QY 241 CTGTCCCAACTCCATTCGCAAGTCCCTGACAGGCTCTCTGACAGCCAGGCTCAAGG 300
DB 1213 CTGTCCCAACTCCATTCGCAAGTCCCTGACAGGCTCTCTGACAGCCAGGCTCAAGG 1272
QY 301 AAGCTCGGAGGCTTTCTCTAGGCTTTTCTATGCGCCCTCCCAAGGACTCTTTCTCTGAA 360
DB 1273 AAGCTCGGAGGCTTTCTCTAGGCTTTTCTATGCGCCCTCCCAAGGACTCTTTCTCTGAA 1332
QY 361 CCCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTATGTTAGGCT 420
DB 1333 CCCTGTAGGCTTTGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTATGTTAGGCT 1392
QY 421 TTTGGTGGAGCCCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGAGAGC 480
DB 1393 TTTGGTGGAGCCCGCAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGAGAGC 1452
QY 481 TTAGCAATCTTTTGCACACT 501
DB 1453 TTAGCAATCTTTTGCACACT 1473

RESULT 9
AAV74190
ID AAV74190 standard; DNA: 2346 BP.
XX AC AAV74190;
XX DT 15-MAR-1999 (first entry)
XX DE Human sgk DNA.
XX KW Serum glucocorticoid regulated kinase; sgk; human; treatment: inhibitor;
XX KW serine/threonine protein kinase family; antagonist; diabetic nephropathy;
XX KW chronic renal failure; inflammation; Alzheimers disease; wound; ss.
XX OS Homo sapiens.
XX Key location/Qualifiers
XX CDS 45..1340
XX FT /*tag- a
XX FT /product- "sgk"
XX FT /transl_except- (POS:1185..1187, aa:Asp)
XX
XX EP887081-A2.
XX 30-DEC-1998.
XX 27-MAY-1998; 98EP-0304189.
XX 27-JUN-1997; 97US-0051124.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Kumar JM;
XX WPI: 1999-047627/05.
XX P-PSDB: AAW90139.
XX
XX Treating chronic renal failure, diabetic nephropathy and Alzheimer's
XX disease - by administration of nucleic acids and antagonists which
XX inhibit activity or expression of human serum glucocorticoid
XX regulated kinase (sgk), a serine/threonine protein kinase
XX
XX Disclosure; Page 14-15; 17pp; English.
XX
XX This sequence encodes a novel human serum glucocorticoid regulated kinase
XX (sgk) protein which is a member of the serine/threonine protein kinase
XX family. This protein is used for the treatment of a subject having need
XX to inhibit/antagonise activity or expression of human sgk polypeptide
XX e.g. for the treatment of chronic renal failure, diabetic nephropathy,
XX inflammation, Alzheimers disease and wounds.
XX
XX Sequence 2346 BP; 608 A; 520 C; 518 G; 700 T; 0 other;
Query Match 99.7%; Score 499.4; DB 20; Length 2346;
Best Local Similarity 99.8%; Pred. No. 1.7e-152;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCTCAGCTGAACCAATATATACAAATTCGGCAAGACACCTCTCGAGGGCTCTCTGC 60
DB 982 CTCTCAGCTGAACCAATATATACAAATTCGGCAAGACACCTCTCGAGGGCTCTCTGC 1041
QY 61 AGAAGACAGGACCAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAGAGTCATG 120
DB 1042 AGAAGACAGGACCAAGCGGCTCGGGCCCAAGGATGACTTTCATGAGATTAAGAGTCATG 1101
QY 121 TCTTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 180
DB 1102 TCTTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1161
QY 181 ACCCAATGTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGC 240
DB 1162 ACCCAATGTGAGTGGGCGCCCAACGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGC 1221

241	CTGTCCCAACTCCATTGGCAAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG	300
1222	CTGTCCCAACTCCATTGGCAAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG	1281
301	AAGCTGCGGAGGCGTTTCCCTAGGCGTTTCTATGCGCCTCCCCACGCGACTCTTCTCTCTGAA	360
1282	AAGCTGCGGAGGCGTTTCCCTAGGCGTTTCTATGCGCCTCCCCACGCGACTCTTCTCTCTGAA	1341
361	CCCTGTTAGGCGTGGTGTTAAGGAAATTATGTCGTGTTCCGAACTGTTTAGTACGCT	420
1342	CCCTGTTAGGCGTGGTGTTAAGGAAATTATGTCGTGTTCCGAACTGTTTAGTACGCT	1401
421	TTTGGTGGAGCCGCCAGCTGACAGGACATCTTACAAAGAGAAATTTGCACATCTCTGGAAGC	480
1402	TTTGGTGGAGCCGCCAGCTGACAGGACATCTTACAAAGAGAAATTTGCACATCTCTGGAAGC	1461
481	TTAGCAATCTTATTGCACACT	501
1462	TTAGCAATCTTATTGCACACT	1482

RESULT 10
AAAF44639
ID AAAF44639 standard: cDNA: 1296 bp.

27-MAR-2001 (first entry)

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
PT treating immune-related diseases and disorders, cardiovascular disease,
PT neurodegenerative diseases and/or cancers -

The present sequence encodes a novel protein kinase. The nucleic acids and the protein kinases they encode may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies and kinase antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune

CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure,
CC oxidative-stress related disorders, chronic inflammatory bowel disease,
CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
CC osteoarthritis, psoriasis, thinitis, autoimmune, diabetes, cancers and
CC reproductive disorders.
XX
XX Sequence 1296 BP: 345 A; 333 C; 293 G; 325 T; 0 other;

RESULT 11
AAF44640
ID AAF44640

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteoblastic;
immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.

OS Mus musculus.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and

PT treating immune-related diseases and disorders, cardiovascular disease,
 XX neurodegenerative diseases and/or cancers -
 PS Disclosure; Fig 2: 310pp; English.
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC or kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.

SQ Sequence 2432 BP; 586 A; 586 C; 560 G; 699 T; 1 other;

Query Match 62.7%; Score 314.2; DB 22; Length 2432;
 Best Local Similarity 81.7%; Pred. No. 7,1e-92;
 Matches 401; Conservative 0; Mismatches 83; Indels 7; Gaps 3;

QY 1 CTCCTCCAGCTGAACCAAAATATTACAAATTCGCCAGACACCTCTCGAGGGCCCTCTGC 60
 DB 1009 CTCCTCCAGTTGAACCAAAATATTACAAATTCGCCAGAGGACCTCTCGAGGGCCCTCTGC 1068
 QY 61 AGAAGGACAGACAAAGCGCTCGGGCCAAAGGATGACCTTCATGGAGATTAAAGATCATG 120
 DB 1069 AGAAGGACCGACCAAGAGGCTGGTGCCAAAGATGACCTTCATGGAGATTAAAGATCATG 1128
 QY 121 TCTTCTCTCTTAACTGAGTGCATCTCTATTAAGAGATTAATCTCCCTCTTTA 180
 DB 1129 TTTTCTCTCTTAACTGAGTGCATCTCTATTAAGAGATTAATCTCCCTCTTTA 1188
 QY 181 ACCCAATGTGAGTGGGCCCAACGAGCTAGCGACCTTTGACCCCGAGTTTACCGAAGAC 240
 DB 1189 ACCCAATGTGAGTGGGCCCAACGAGCTAGCGACCTTCGATCCCGAGTTTACCGAAGAC 1248
 QY 241 CTGTCCCAACATCATTCATGGCAAGTCCCTCGACAGGCTCTCGTCACAGCAGCGTCAAGG 300
 DB 1249 CGGTCCCAACATCATTCATGGCAAGTCCCTCGACAGCAGCTCTGTGACGGCCAGTGTGAAG 1308
 QY 301 AAGCTGCGAGGCTTCTCTAGGCTTTTCTCTATGCGCTCCCGACGGACTCTTCTCTGAA 360
 DB 1309 AAGCAGCAGAGCCCTTCTCGCTTCTCTATGCACTCTCTGTGGATTCTCTCTCTGAG 1368
 QY 361 CCCGTGAGGCTTGGTTTAAAGGATTTATGTGTGTTCCGANTGT-TTATGTTAGGC 419
 DB 1369 TGC--TCCCGGAGTGGTCTGAAGGACTTCTCGAGGCTTCTCTGAGGCTTCTCTGTTAGC 1426
 QY 420 TTTTGGTGGAGCCCGAGCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTCGAAG 479
 DB 1427 CTTTGGTGGAGTTCGAGCTGACAGACATTTTAAAGAAATTTGCACAC---CTGGAAG 1482
 QY 480 CTTAGCAATCT 490
 DB 1483 CTTGCACTCT 1493
 RESULT 12
 AAA46592
 ID AAA46592 standard; DNA; 2365 BP.
 XX
 AC AAA46592;
 XX
 DT 25-SEP-2000 (first entry)

XX DNA encoding a rat serum and glucocorticoid induced protein kinase.
 DE
 XX Protein kinase; Pkh1; Pkh2; Ypk1; Yrk2; protein kinase B-alpha;
 KW serum and glucocorticoid induced protein kinase; SGK; PKAalpa;
 KW 3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
 KW thrush; cancer; diabetes; obesity; antifungal; Candida infection; ss.
 XX
 OS Rattus sp.
 PH key Location/Qualifiers
 FT 1..1288
 CDS /tag- a
 FT /product- "serum and glucocorticoid induced protein
 FT kinase (SGK)"
 FT /trans_except- (pos: 1, aa: Thr)
 FT /note- "the codon encoding Met at position 1 and
 FT Pro at position 318 are not given"
 XX
 PN W0200036135-A2.
 XX
 PD 22-JUN-2000.
 XX
 PF 14-DEC-1999; 99WO-GB04228.
 XX
 PR 14-DEC-1998; 98US-0112114.
 XX
 PA (MEDI-) MEDICAL RES COUNCIL.
 PA (RECC) UNIV CALIFORNIA.
 XX
 XX Thorner JW, Alessi DR, Torrance PD, Casamayor A;
 PI WPI; 2000-442391/38.
 DR P-PSDB; AAY93530.
 XX
 XX Screening method identifying compounds which modulate protein kinase
 PT activity for use in treating fungal infections and cancer -
 PT
 XX Disclosure; Fig 12; 155pp; English.
 XX
 CC The present sequence encodes a rat serum and glucocorticoid induced
 CC protein kinase (SGK). The specification describes a screening method
 CC to identify a compound which modulates the activity of protein kinases
 CC from different sources using host yeast cells. The method is used to
 CC identify a compound which modulates (inhibits) the activity of a
 CC protein kinase. Pkh1 or Pkh2 phosphorylate and activate Ypk1, Yrk2,
 CC SGK or protein kinase B-alpha (PKAalpa). 3-phosphoinositide-dependent
 CC protein kinase-1 (PDK1) is used to phosphorylate and activate Ypk1 and
 CC Yrk2 or SGK but not PKAalpa or p70S6 kinase. Compounds identified by
 CC the methods are used to treat fungal infections e.g. thrush, and to
 CC treat cancer. To treat cancer, the compounds which activate PKB or PDK1 or the
 CC activation of PKB by PDK1. Compounds which activate PKB or PDK1 can be
 CC used in the treatment of diabetes or obesity, and compounds which
 CC inhibit a fungal functional homologue of PDK1 (Pkh1 or Pkh2) or SGK
 CC infections, e.g. thrush.
 CC
 SQ Sequence 2365 BP; 591 A; 580 C; 518 G; 676 T; 0 other;
 Query Match 59.8%; Score 299.4; DB 21; Length 2365;
 Best Local Similarity 78.4%; Pred. No. 4.9e-87;
 Matches 371; Conservative 0; Mismatches 101; Indels 1; Gaps 1;
 QY 17 AAATATTACAAATTCGCCAGACACCTCTCGAGGGCCCTCTCGAGAGGACGAGCAAA 76
 DB 946 AAATATACCAACTCAGCAGGACCTCTCGAGGGCCCTCTCGAGAGGACGAGCAAA 1005
 QY 77 CGCGCTCGGGCCGAGGATGACTCTCTGAGATTAGATCTCTTCTCTCTTAT 136
 DB 1006 GAGCGCTGGTCCCAAGGATGACTTTATGAGATTAGATCTCTTCTCTTAT 1065
 QY 137 TAAGTGGATGATCTCTTATTAAGAAGATTACTCCCTCTTTAAACCCAAATCTAGTGG 196
 |||||||

DB 1066 TAAGTGGATGATCTCATTAATAAGAGATCAGCCGCCCATTTAAACCCAAATCTGAGCGG 1125
 QY 197 GCCAAGAGCTAGGACCTTTGACCCGAGTTTACCGAAGAGCGCTGTCCTCCCACTCAT 256
 DB 1126 GCCAGTACCTTCGGGACCTTGTATCCGAGTTTACTGAGAGCGGTCCTCCAGCTCCAT 1185
 QY 257 TGGCAAGTCCCTGACAGGCTCTGTCACAGCCGCTCAAGGAAGCTGCGGAGGCTTT 316
 DB 1186 CGGCGATCCCTGACAGCATCTTGTTCACAGCCAGTGTGAAGAAGCGCGGAAGCCTT 1245
 QY 317 CTAGGCTTTCTCTATCGCTCCGAGGACTCTTCTCTGAACCTCTGTAGGGCTTGG 376
 DB 1246 CTGTGGCTTCCTATGCGCTCTATGAGACTCTTCTCTGAAGCTCCGAGGTGTT 1305
 QY 377 TTTTAAAGATTTATGTGTGTTTCCGAAATGTTTGTAGTACGCTTTTGTGAGCGGCA 436
 DB 1306 CGAAGGATTTCTCCACGCTTTTCTAAAGTGTGTTTGTAGTACGCTTTGTTGGAGTTACCA 1365
 QY 437 GCTGACAGGACATCTTACAGAGAAATTTGCACATCTCTGGAAGCTTAGCAATC 489
 DB 1366 GCTGACAGGACATCTTAGAAGAGAAATTTGCACAGCAGGAAGCTTGGCAGTC 1417
 RESULT 13
 ABX38963
 ID ABX38963 standard; cDNA; 433 BP.
 XX
 AC ABX38963;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #4128.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding;
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-0960352.
 XX
 PR 12-JAN-1999; 99US-115707P.
 PR 11-JAN-2000; 2000US-0480902.
 XX
 PA (BYATT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARREN/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and
 PT analysis, cattle breeding, or for genetically improving cattle
 XX
 PS Claim 2; SEQ ID NO 4128; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived to a
 CC from cattle, and the LMFD nucleic acid can specifically hybridise to a
 CC second nucleic acid molecule comprising any of 1512 nucleotide
 CC sequences, appearing as ABX34836-ABX4947, or complements of them.
 CC Also included are: (1) a transformed cell having a nucleic acid
 CC comprising an LMFD nucleic acid linked to a promoter and a 3' non-
 CC translated sequence that functions in the cell to cause termination of
 CC transcription and addition of polyadenylated ribonucleotides to a 3' end
 CC of the mRNA molecule; and (2) determining a level or pattern of a

CC molecule in a bovine cell or tissue comprising: (a) incubating a marker
 CC nucleic acid (comprising any of the 1512 nucleic acid sequences of its
 CC complement or fragment) with a complementary nucleic acid molecule
 CC obtained from the bovine cell or tissue, where hybridisation between the
 CC marker nucleic acid and the complementary nucleic acid permits the
 CC detection of the molecule; and (b) detecting the level or pattern of the
 CC complementary nucleic acid, where the detection of the complementary
 CC nucleic acid is predictive of the level or pattern of the molecule.
 CC The LMFD nucleic acid is used for determining a level or pattern
 CC of a molecule in a bovine cell or tissue. It is useful for genome
 CC mapping, gene identification and analysis, cattle breeding, preparation
 CC of constructs for use in cattle gene expression, or for genetically
 CC improving cattle. The present sequence is one of the 1512 bovine
 CC LMFD EST (expressed sequence tag) nucleic acids.
 CC Note: The present sequence was not shown in the specification but
 CC was obtained in electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139.

XX SQ Sequence 433 BP; 107 A; 115 C; 105 G; 106 T; 0 other;

Query Match 58.6%; Score 293.4; DB 25; Length 433;
 Best Local Similarity 83.5%; Pred. No. 1.8e-85;
 Matches 365; Conservative 0; Mismatches 51; Indels 21; Gaps 2;

QY 47 GGAGGGCTCTCGCAAGAGACAGACAGCGCGGCGGCAAGATGATTCATGCA 106
 DB 12 GGAAGGGCTCTCGCAAGAGACAGACAGCGCGGCGGCAAGATGATTCATGCA 71
 QY 107 GATTAAGAGTCATGCTTCTCTCTTAATTAATTAATTAATTAATTAATTA 166
 DB 72 GATTAAGAGTCATGCTTCTCTCTTAATTAATTAATTAATTAATTAATTA 141
 QY 167 TACTCCCTCTTTAAACCAATGAGTGGGCGGCGGCGGCGGCGGCGGCGGCG 226
 DB 132 TACTCCCTCTTTAAACCAATGAGTGGGCGGCGGCGGCGGCGGCGGCGGCG 191
 QY 227 GTTACCGAAGAGCTGTCGCGCACTCCCTTGAAGTGGGCGGCGGCGGCGGCG 286
 DB 192 GTTACCGAAGAGCTGTCGCGCACTCCCTTGAAGTGGGCGGCGGCGGCGGCG 251
 QY 287 AGCCAGGCTCAAGAGCTGCGGAGCTTCTTCTAGGCTTTCTATGCGGCGGCGG 345
 DB 252 AGCCAGGCTCAAGAGCTGCGGAGCTTCTTCTAGGCTTTCTATGCGGCGGCGG 411
 QY 346 ACTCTTCTCTGAAACCTGTTAGGCTTCTTGAAGGATTTTATGCTGTCTGCA 405
 DB 312 ACTCTTCTCTGAAACCTGTTAGGCTTCTTGAAGGATTTTATGCTGTCTGCA 351
 QY 406 TGTCTTCTAGGCTTTTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 465
 DB 352 TTTTCTTCTAGGCTTTTGTGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 411
 QY 466 CACATCTCTGGAAGCTT 482
 DB 412 CACATCTCTGGAAGCTT 428

RESULT 14
 ABX42261
 ID ABX42261 standard; cDNA; 404 BP.
 XX
 AC ABX42261;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #7426.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding;
 XX
 OS Bos Taurus.
 XX

QY 308 CGAGGCTTCCCTAGGCTTTTCCTATGCGCCTCCACGGACTCTTTCCCTCTGAACCTGTT 367
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
111 CGAGGCTTCCCTAGGCTTTTCCTATGCGCCTCCACGGACTCTTTCCCTCTGAACCTGTT 170
QY 368 AGGGCTTGCTTTTAAAGGATTTATGCTGTGTTCCGAATGTTTCTAGTACCTTTGGTG 427
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
171 AGGGCTTGCTTTTAAAGGATTTATGCTGTGTTCCGAATGTTTCTAGTACCTTTGGTG 230
QY 428 GAGCCGCCAGCTGACAGGACATCTTACAAGGAATTTGCACATCTCTGGAAGCTTAGCAA 487
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
231 GAGCCGCCAGCTGACAGGACATCTTACAAGGAATTTGCACATCTCTGGAAGCTTAGCAA 290
QY 488 TCTTATTGCACACT 501
Db ||||||||||||||||
291 TCTTATTGCACACT 304

Search completed: August 11, 2003, 13:51:06
Job time : 253 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 13:40:45 ; Search time 2064 Seconds
(without alignments)
5899,488 Million cell updates/sec

Title: us-10-000-039a-1_copy_980_1480

Perfect score: 501
Sequence: 1 CTTCCAGCTGAACCAAT.....TAGCAATCTATTGCACACT 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estm.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	499.4	99.7	608	14	CB216184
2	499.4	99.7	630	14	CB216592
3	499.4	99.7	671	12	BC770190
4	499.4	99.7	765	12	BC763506

5	499.4	99.7	859	13	BC680126
6	499.4	99.7	873	10	BC756781
7	499.4	99.7	912	13	BC756781
8	499.4	99.7	912	13	BC756781
9	499.4	99.7	912	13	BC756781
10	499.4	99.7	912	13	BC756781
11	499.4	99.7	912	13	BC756781
12	499.4	99.7	912	13	BC756781
13	499.4	99.7	912	13	BC756781
14	499.4	99.7	912	13	BC756781
15	499.4	99.7	912	13	BC756781
16	499.4	99.7	912	13	BC756781
17	499.4	99.7	912	13	BC756781
18	499.4	99.7	912	13	BC756781
19	499.4	99.7	912	13	BC756781
20	499.4	99.7	912	13	BC756781
21	499.4	99.7	912	13	BC756781
22	499.4	99.7	912	13	BC756781
23	499.4	99.7	912	13	BC756781
24	499.4	99.7	912	13	BC756781
25	499.4	99.7	912	13	BC756781
26	499.4	99.7	912	13	BC756781
27	499.4	99.7	912	13	BC756781
28	499.4	99.7	912	13	BC756781
29	499.4	99.7	912	13	BC756781
30	499.4	99.7	912	13	BC756781
31	499.4	99.7	912	13	BC756781
32	499.4	99.7	912	13	BC756781
33	499.4	99.7	912	13	BC756781
34	499.4	99.7	912	13	BC756781
35	499.4	99.7	912	13	BC756781
36	499.4	99.7	912	13	BC756781
37	499.4	99.7	912	13	BC756781
38	499.4	99.7	912	13	BC756781
39	499.4	99.7	912	13	BC756781
40	499.4	99.7	912	13	BC756781
41	499.4	99.7	912	13	BC756781
42	499.4	99.7	912	13	BC756781
43	499.4	99.7	912	13	BC756781
44	499.4	99.7	912	13	BC756781
45	499.4	99.7	912	13	BC756781

ALIGNMENTS

RESULT 1	CB216184	NISC_nq02c07.y1	NICHD_HS_U2c	Homo sapiens	cdna	clone	linear	EST 06-FEB-2003
LOCUS	CB216184	5' mRNA sequence						5937781
DEFINITION	CB216184	5' mRNA sequence						
ACCESSION	CB216184	GI:28264376						
VERSION	CB216184	EST						
KEYWORDS	CB216184	EST						
SOURCE	CB216184	EST						
ORGANISM	CB216184	EST						
REFERENCE	CB216184	EST						
AUTHORS	CB216184	EST						
TITLE	CB216184	EST						
JOURNAL	CB216184	EST						
COMMENT	CB216184	EST						

Plate: LLM13165 row: E column: 14

Seq primer: M13RPI reverse primer (ABI).

FEATURES
Source
1. 608
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5937781"
/sex="female"
/tissue_type="normal endometrium, mid-secretory phase,
cycle day 23"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NICHD_HS_Ut2"
/note="Organ: uterus; Vector: pCMV-SPORT6.1.cdb (ResGen,
Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
Cloned unidirectionally from microquantity amounts of mRNA
from normal endometrial tissue (mid-secretory phase, cycle
day 23). Average insert size 1.6 kb. Library constructed
by ResGen (Invitrogen Corporation)."

BASE COUNT 151 a 184 c 132 g 161 t

Query Match 99.7%; Score 499.4; DB 14; Length 608;
Best Local Similarity 99.8%; Pred. No. 2.5e-134;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTCTCCAGCTGAACCAATATTACAAATTCGCCACAGACACCTCTCTGGAGGCTCTCTGC 60
Db 60 CTCTCCAGCTGAACCAATATTACAAATTCGCCACAGACACCTCTCTGGAGGCTCTCTGC 119
Qy 61 AGAAGACAGACAGCAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGATCATG 120
Db 120 AGAAGACAGACAGCAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGATCATG 179
Qy 121 TCTTCTCTCTTAATTAATCTGGGATGATCTCAATTAAGAAGATTACTCCGCCCTTTTA 180
Db 180 TCTTCTCTCTTAATTAATCTGGGATGATCTCAATTAAGAAGATTACTCCGCCCTTTTA 239
Qy 181 ACCCAATGAGTGGGCCCAACGAGCTACGGACCTTTGACCGGAGTTTACCGAGAGC 240
Db 240 ACCCAATGAGTGGGCCCAACGAGCTACGGACCTTTGACCGGAGTTTACCGAGAGC 299
Qy 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGAGGCTCAAGG 300
Db 300 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGAGGCTCAAGG 359
Qy 301 AAGTCGCGAGGCTTCTAGGCTTTTCCTATGCGCTCCGAGGACTCTTCTCTGAA 360
Db 360 AAGTCGCGAGGCTTCTAGGCTTTTCCTATGCGCTCCGAGGACTCTTCTCTGAA 419
Qy 361 CCCTGTAGGCTTGGTTTTAAAGGATTTTATGTGTTTCCGAATTTTATGTTAGTACCT 420
Db 420 CCCTGTAGGCTTGGTTTTAAAGGATTTTATGTGTTTCCGAATTTTATGTTAGTACCT 479
Qy 421 TTGGTGGAGCGCGAGCTGACAGGACATCTTACAGAGAATTGGACATCTCTGGAAGC 480
Db 480 TTGGTGGAGCGCGAGCTGACAGGACATCTTACAGAGAATTGGACATCTCTGGAAGC 539
Qy 481 TTAGCAATCTATTGGACACT 501
Db 540 TTAGCAATCTATTGGACACT 560

RESULT 2
CB216592
LOCUS CB216592 630 bp mRNA linear EST 06-FEB-2003
DEFINITION NISC_bq06g10.y1 NICHD_HS_Ut2 Homo sapiens CDNA clone IMAGE:5938363
5', mRNA sequence.
ACCESSION CB216592
VERSION CB216592.1 GI:28264784
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 630)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapb-r@mail.nih.gov
CDNA Library Preparation:
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL
Sequencing Center (NISC)
Cloned unidirectionally from microquantity amounts of mRNA
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Plate: LLML3166 row: M column: 20
Seq primer: M13RPI reverse primer (ABI).

FEATURES
Source

1. 630
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5938363"
/sex="female"
/tissue_type="normal endometrium, mid-secretory phase,
cycle day 23"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NICHD_HS_Ut2"
/note="Organ: uterus; Vector: pCMV-SPORT6.1.cdb (ResGen,
Invitrogen Corporation); Site_1: NotI; Site_2: EcoRV;
Cloned unidirectionally from microquantity amounts of mRNA
from normal endometrial tissue (mid-secretory phase, cycle
day 23). Average insert size 1.6 kb. Library constructed
by ResGen (Invitrogen Corporation)."

BASE COUNT 153 a 167 c 141 g 169 t

Query Match 99.7%; Score 499.4; DB 14; Length 630;
Best Local Similarity 99.8%; Pred. No. 2.5e-134;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTCTCCAGCTGAACCAATATTACAAATTCGCCACAGACACCTCTCTGGAGGCTCTCTGC 60
Db 94 CTCTCCAGCTGAACCAATATTACAAATTCGCCACAGACACCTCTCTGGAGGCTCTCTGC 153
Qy 61 AGAAGACAGACAGCAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGATCATG 120
Db 154 AGAAGACAGACAGCAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGATCATG 213
Qy 121 TCTTCTCTCTTAATTAATCTGGGATGATCTCAATTAAGAAGATTACTCCGCCCTTTTA 180
Db 214 TCTTCTCTCTTAATTAATCTGGGATGATCTCAATTAAGAAGATTACTCCGCCCTTTTA 273
Qy 181 ACCCAATGAGTGGGCCCAACGAGCTACGGACCTTTGACCGGAGTTTACCGAGAGC 240
Db 274 ACCCAATGAGTGGGCCCAACGAGCTACGGACCTTTGACCGGAGTTTACCGAGAGC 333
Qy 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGAGGCTCAAGG 300
Db 334 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGAGGCTCAAGG 393
Qy 301 AAGTCGCGAGGCTTCTAGGCTTTTCCATGCGCTCCGAGGACTCTTCTCTGAA 360
Db 394 AAGTCGCGAGGCTTCTAGGCTTTTCCATGCGCTCCGAGGACTCTTCTCTGAA 453
Qy 361 CCCTGTAGGCTTGGTTTTAAAGGATTTTATGTGTTTCCGAATTTTATGTTAGTACCT 420
Db 454 CCCTGTAGGCTTGGTTTTAAAGGATTTTATGTGTTTCCGAATTTTATGTTAGTACCT 513
Qy 421 TTGGTGGAGCGCGAGCTGACAGGACATCTTACAGAGAATTGGACATCTCTGGAAGC 480
Db 514 TTGGTGGAGCGCGAGCTGACAGGACATCTTACAGAGAATTGGACATCTCTGGAAGC 573

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCTGGAGGGCTCTCTGC 60
 DB 242 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCTGGAGGGCTCTCTGC 301
 QY 61 AGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGAGTCATG 120
 DB 302 AGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGAGTCATG 361
 QY 121 TCTTCT 180
 DB 362 TCTTCT 421
 QY 181 ACCCAATGTAGTGGGCCCAAGCAGCTAGGCACTTTGACCCGAGCTTTACCGAAGC 240
 DB 422 ACCCAATGTAGTGGGCCCAAGCAGCTAGGCACTTTGACCCGAGCTTTACCGAAGC 481
 QY 241 CTGTCCCACTCTCATTTGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCCTCAAGG 300
 DB 482 CTGTCCCACTCTCATTTGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCCTCAAGG 541
 QY 301 AAGTGGCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCT 360
 DB 542 AAGTGGCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCT 601
 QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTATGTTGTTCCGAGTCTTTAGTAGGCT 420
 DB 602 CCCTGTAGGCTTGGTTTAAAGGATTTATGTTGTTCCGAGTCTTTAGTAGGCT 661
 QY 421 TTGTGGAGCGCGCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGGAAGC 480
 DB 662 TTGTGGAGCGCGCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGGAAGC 721
 QY 481 TTAGCAATCTTATTGCACACT 501
 DB 722 TTAGCAATCTTATTGCACACT 742

RESULT 5

B0680126
 LOCUS
 DEFINITION B0680126 859 bp mRNA linear EST 15-JUL-2002
 5', mRNA sequence.

ACCESSION B0680126
 VERSION B0680126.1 GI:21792805

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 859)
 NIH-MGC http://mgi.nci.nih.gov/

AUTHORS
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished

COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov

Tissue Procurement: DCTD/DTF
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: L1CM2427 row: 1 column: 19
 High quality sequence stop: 718.

FEATURES
 Location/Qualifiers

1..859
 /organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6263586"

/tissue_type="melanotic melanoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH-MGC_112"

/note="Organ: skin; Vector: pOTB7; Site: 1; XhoI: Site-2;
 EcoRI: cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(S). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH-MGC Library."

BASE COUNT 206 a 211 c 201 g 241 t
 ORIGIN

Query Match 99.7%; Score 499.4; DB 13; Length 859;
 Best Local Similarity 99.8%; Pred. No. 2.8e-134;
 Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCTGGAGGGCTCTCTGC 60
 DB 96 CTCTCCAGCTGAACCAAAATATTACAAATTCGCAAGACACACCTCTCTGGAGGGCTCTCTGC 155
 QY 61 AGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGAGTCATG 120
 DB 156 AGAAGGACAGACAAACCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAAGAGTCATG 215
 QY 121 TCTTCT 180
 DB 216 TCTTCT 275
 QY 181 ACCCAATGTAGTGGGCCCAAGCAGCTAGGCACTTTGACCCGAGCTTTACCGAAGC 240
 DB 276 ACCCAATGTAGTGGGCCCAAGCAGCTAGGCACTTTGACCCGAGCTTTACCGAAGC 335
 QY 241 CTGTCCCACTCTCATTTGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCCTCAAGG 300
 DB 336 CTGTCCCACTCTCATTTGCAAGTCCCTGACAGCGTCTCTGACAGCCAGCCTCAAGG 395
 QY 301 AAGTGGCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCT 360
 DB 396 AAGTGGCGAGGCTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCTTAGGCTTTCT 455
 QY 361 CCCTGTAGGCTTGGTTTAAAGGATTTATGTTGTTCCGAGTCTTTAGTAGGCT 420
 DB 456 CCCTGTAGGCTTGGTTTAAAGGATTTATGTTGTTCCGAGTCTTTAGTAGGCT 515
 QY 421 TTGTGGAGCGCGCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGGAAGC 480
 DB 516 TTGTGGAGCGCGCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTGGAAGC 575
 QY 481 TTAGCAATCTTATTGCACACT 501
 DB 576 TTAGCAATCTTATTGCACACT 596

RESULT 6

BG756781
 LOCUS
 DEFINITION BG756781 873 bp mRNA linear EST 15-MAY-2001
 602710254F1 NIH-MGC_48 Homo sapiens cDNA clone IMAGE:4850820 5',
 mRNA sequence.

ACCESSION BG756781
 VERSION BG756781.1 GI:14067434

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 873)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1692 row: k column: 13
High quality sequence stop: 797.
Location/Qualifiers
1. 873

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4850820"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_48"
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 201 a 211 c 205 g 256 t

Query Match 99.7%; Score 499.4; DB 10; Length 873;
Best Local Similarity 99.8%; Pred. No. 2.8e-134;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCCAGACACCTCTGGAGGGCTCTCTGC 60
Db 4 CTCCTCAGCTGAACCAATATTACAAATTCGCCAGACACCTCTGGAGGGCTCTCTGC 63
QY 61 AGAAGACAGGACAAACGGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 120
Db 64 AGAAGACAGGACAAACGGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 123
QY 121 TCCTCTCTCCTTAATTAAGTGGATGATCTCATTATTAAGAAGATTACTCCCTCTTTA 180
Db 124 TCCTCTCTCCTTAATTAAGTGGATGATCTCATTATTAAGAAGATTACTCCCTCTTTA 183
QY 181 ACCCAATGTAGTGGGCCCAACAGCTAGCGACCTTTGACCCCGAGTTTACCGAGAGC 240
Db 184 ACCCAATGTAGTGGGCCCAACAGCTAGCGACCTTTGACCCCGAGTTTACCGAGAGC 243
QY 241 CTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGCTCAAGG 300
Db 244 CTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGCTCAAGG 303
QY 301 AAGCTGCGGAGGCTTTCTAGGCTTTTCTATGGGCTCCACGGAGCTTTCTCTTGAA 360
Db 304 AAGCTGCGGAGGCTTTCTAGGCTTTTCTATGGGCTCCACGGAGCTTTCTCTTGAA 363
QY 361 CCCTGTTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATCTTTTAGTAGCCCT 420
Db 364 CCCTGTTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATCTTTTAGTAGCCCT 423
QY 421 TTTGCTGGAGCCGACGCTGACAGGACATCTTACAAGAGAATTTCACATCTCTGGAAGC 480
Db 424 TTTGCTGGAGCCGACGCTGACAGGACATCTTACAAGAGAATTTCACATCTCTGGAAGC 483
QY 481 TTAGCAATCTTATTGCACACT 501
Db 484 TTAGCAATCTTATTGCACACT 504

RESULT 7
BQ676527 912 bp mRNA linear EST 15-JUL-2002
LOCUS BQ676527
DEFINITION AGENCOURT_8196995 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6263470
5', mRNA sequence.

ACCESSION BQ676527
VERSION BQ676527.1
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 912)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-rmail.nih.gov
Tissue Procurement: DCRD/DFP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2427 row: g column: 23
High quality sequence stop: 678.
Location/Qualifiers
1. 912

FEATURES
Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6263470"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 218 a 227 c 213 g 254 t

Query Match 99.7%; Score 499.4; DB 13; Length 912;
Best Local Similarity 99.8%; Pred. No. 2.9e-134;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCCAGACACCTCTGGAGGGCTCTCTGC 60
Db 96 CTCCTCAGCTGAACCAATATTACAAATTCGCCAGACACCTCTGGAGGGCTCTCTGC 155
QY 61 AGAAGACAGGACAAACGGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 120
Db 156 AGAAGACAGGACAAACGGCTCGGGCCCAAGATGACTTCATGGAGATTAAAGATCATG 215
QY 121 TCCTCTCTCCTTAATTAAGTGGATGATCTCATTATTAAGAAGATTACTCCCTCTTTA 180
Db 216 TCCTCTCTCCTTAATTAAGTGGATGATCTCATTATTAAGAAGATTACTCCCTCTTTA 275
QY 181 ACCCAATGTAGTGGGCCCAACAGCTAGCGACCTTTGACCCCGAGTTTACCGAGAGC 240
Db 276 ACCCAATGTAGTGGGCCCAACAGCTAGCGACCTTTGACCCCGAGTTTACCGAGAGC 335
QY 241 CTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGCTCAAGG 300
Db 336 CTGTCCCACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCGCTCAAGG 395
QY 301 AAGCTGCGGAGGCTTTCTAGGCTTTTCTATGGGCTCCACGGAGCTTTCTCTTGAA 360
Db 396 AAGCTGCGGAGGCTTTCTAGGCTTTTCTATGGGCTCCACGGAGCTTTCTCTTGAA 455
QY 361 CCCTGTTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATCTTTTAGTAGCCCT 420
Db 456 CCCTGTTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATCTTTTAGTAGCCCT 515

QY 421 TTTCGTGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTCGAAGC 480
 Db 516 TTTCGTGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTCGAAGC 575

QY 481 TTAGCAATCTTATTGCACACT 501
 Db 576 TTAGCAATCTTATTGCACACT 596

RESULT 8
 B0838753 918 bp mRNA linear EST 16-OCT-2002
 LOCUS B0838753.1 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6258012
 DEFINITION 5', mRNA sequence.

ACCESSION B0838753
 VERSION B0838753.1 GI:24023148
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 918)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: DCTD/DTp
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2413 row: d column: 13
 High quality sequence stop: 524.
 Location/Qualifiers
 1..918
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6258012"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="NIH_MGC_112"
 /note="Organ: skin; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

BASE COUNT 214 a 289 c 218 g 197 t
 ORIGIN

Query Match 98.7%; Score 494.6; DB 13; Length 918;
 Best Local Similarity 99.2%; Pred. No. 7.2e-133;
 Matches 497; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAATATTACAAATTTCCCAAGACACCTCTCGGAGGCTCTCTGC 60
 Db 96 CTCTCCAGCTGAACCAATATTACAAATTTCCCAAGACACCTCTCGGAGGCTCTCTGC 155

QY 61 AGAAGACAGACAAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTCATG 120
 Db 156 AGAAGACAGACAAAGCGCTCGGGCCCAAGGATGACTTCATGGAGATTAAGAGTCATG 215

QY 121 TCCTCTCTCTTAATTAAGTGGATGATCTCATTAATACAGATTAATCCCTTTTA 180
 Db 216 TCCTCTCTCTTAATTAAGTGGATGATCTCATTAATACAGATTAATCCCTTTTA 275

QY 181 ACCCAATGTGAGTGGGCCCAACAGAGCTAGCGACCTTTGACCCCGAGCTTTACCGAAGC 240
 Db 276 ACCCAATGTGAGTGGGCCCAACAGAGCTAGCGACCTTTGACCCCGAGCTTTACCGAAGC 315

QY 241 CTGTCCCAACTCCATTTGCAAGTCCCTGACAGCGTCTCGTCACAGCCAGGCTCAAGG 300
 Db 336 CTGTCCCAACTCCATTTGCAAGTCCCTGACAGCGTCTCGTCACAGCCAGGCTCAAGG 395

QY 301 AAGCTGCCGAGGCTTTCTCTAGAGCTTTCTATGCGCTCCACAGGACTCTTTCTCTGAA 360
 Db 396 AAGCTGCCGAGGCTTTCTCTAGAGCTTTCTATGCGCTCCACAGGACTCTTTCTCTGAA 455

QY 361 CCCTGTAGGCGCTTGGTTTAAAGCATTTATGTGTTTCCGGAATCTTTAGTAACT 420
 Db 456 CCCTGTAGGCGCTTGGTTTAAAGCATTTATGTGTTTCCGGAATCTTTAGTAACT 515

QY 421 TTTCGTGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTCGAAGC 480
 Db 516 TTTCGTGAGCGCCAGCTGACAGACATCTTACAAGAGAAATTTGCACATCTCTCGAAGC 575

QY 481 TTAGCAATCTTATTGCACACT 501
 Db 576 TTAGCAATCTTATTGCACACT 596

RESULT 9
 B0769533 899 bp mRNA linear EST 15-MAY-2001
 LOCUS B0769533.1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4877596 5',
 DEFINITION mRNA sequence.

ACCESSION B0769533
 VERSION B0769533.1 GI:14080186
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 899)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: ATCC/BCVD/DTp
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1762 row: q column: 05
 High quality sequence stop: 822.
 Location/Qualifiers
 1..899
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4877596"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="NIH_MGC_49"
 /note="Organ: skin; Vector: pOTB7; Site: 1; XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library. I."

BASE COUNT 203 a 216 c 216 g 264 t
 ORIGIN

FEATURES
 source

```
Query Match      98.3%; Score 492.4; DB 12; Length 899;
Best Local Similarity 99.8%; Pred. No. 3.1e-132;
Matches 493; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCTGAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGGCTCTCTGCAAGAGGA 67
   |||
Db 1 GCTGAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGGCTCTCTGCAAGAGGA 60
   |||
QY 68 CAGACAAGGGCTCGGGGCAAGGATGACTTCATGGAGATTAAAGAGTCAATGCTCTT 127
   |||
Db 61 CAGACAAGGGCTCGGGGCAAGGATGACTTCATGGAGATTAAAGAGTCAATGCTCTT 120
   |||
QY 128 CTCCTTAATTAACGGGATGATCTCATTAAAGAAGATTAATCCCTTTTAAACCAAA 187
   |||
Db 121 CTCCTTAATTAACGGGATGATCTCATTAAAGAAGATTAATCCCTTTTAAACCAAA 180
   |||
QY 188 TGTGAGTGGGCGCCCAAGCTTACGCACTTTGACCCGAGTTTACCGAAGAGCTGTCC 247
   |||
Db 181 TGTGAGTGGGCGCCCAAGCTTACGCACTTTGACCCGAGTTTACCGAAGAGCTGTCC 240
   |||
QY 248 CAATCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGGAAGTGC 307
   |||
Db 241 CAATCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGGAAGTGC 300
   |||
QY 308 CGAGGCTTCTAGGCTTTTCTATGCGCTTCCACAGGACCTCTCTCTGCAAGAGTGC 367
   |||
Db 301 CGAGGCTTCTAGGCTTTTCTATGCGCTTCCACAGGACCTCTCTCTGCAAGAGTGC 360
   |||
QY 368 AGGCTTGGTTTTAAAGCAATTTATGTGTGTTCCGAATGTTTATGTTAGTCTTCTG 427
   |||
Db 361 AGGCTTGGTTTTAAAGCAATTTATGTGTGTTCCGAATGTTTATGTTAGTCTTCTG 420
   |||
QY 428 GAGCGCCAGCTGACAGGACATCTTACAGAAGATTTCACATCTCTGGAAGCTTAGAA 487
   |||
Db 421 GAGCGCCAGCTGACAGGACATCTTACAGAAGATTTCACATCTCTGGAAGCTTAGAA 480
   |||
QY 488 TCTTATTCACACT 501
   |||
Db 481 TCTTATTCACACT 494
   |||

RESULT 11
LOCUS      CB154488
DEFINITION K-EST0212421 B2N807043 Homo sapiens CDNA clone B2N807043-27-D05 5',
            mRNA sequence.
ACCESSION  CB154488
VERSION     CB154488.1 GI:28139489
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 561)
AUTHORS     Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
            Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
            Kim, Y.S.

Query Match      98.3%; Score 492.4; DB 12; Length 899;
Best Local Similarity 99.8%; Pred. No. 3.1e-132;
Matches 493; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 GCTGAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGGCTCTCTGCAAGAGGA 67
   |||
Db 1 GCTGAACCAATATTACAAATTCGCAAGACACCTCTCTGGAGGGCTCTCTGCAAGAGGA 60
   |||
QY 68 CAGACAAGGGCTCGGGGCAAGGATGACTTCATGGAGATTAAAGAGTCAATGCTCTT 127
   |||
Db 61 CAGACAAGGGCTCGGGGCAAGGATGACTTCATGGAGATTAAAGAGTCAATGCTCTT 120
   |||
QY 128 CTCCTTAATTAACGGGATGATCTCATTAAAGAAGATTAATCCCTTTTAAACCAAA 187
   |||
Db 121 CTCCTTAATTAACGGGATGATCTCATTAAAGAAGATTAATCCCTTTTAAACCAAA 180
   |||
QY 188 TGTGAGTGGGCGCCCAAGCTTACGCACTTTGACCCGAGTTTACCGAAGAGCTGTCC 247
   |||
Db 181 TGTGAGTGGGCGCCCAAGCTTACGCACTTTGACCCGAGTTTACCGAAGAGCTGTCC 240
   |||
QY 248 CAATCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGGAAGTGC 307
   |||
Db 241 CAATCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGGAAGTGC 300
   |||
QY 308 CGAGGCTTCTAGGCTTTTCTATGCGCTTCCACAGGACCTCTCTCTGCAAGAGTGC 367
   |||
Db 301 CGAGGCTTCTAGGCTTTTCTATGCGCTTCCACAGGACCTCTCTCTGCAAGAGTGC 360
   |||
QY 368 AGGCTTGGTTTTAAAGCAATTTATGTGTGTTCCGAATGTTTATGTTAGTCTTCTG 427
   |||
Db 361 AGGCTTGGTTTTAAAGCAATTTATGTGTGTTCCGAATGTTTATGTTAGTCTTCTG 420
   |||
QY 428 GAGCGCCAGCTGACAGGACATCTTACAGAAGATTTCACATCTCTGGAAGCTTAGCAA 487
   |||
Db 421 GAGCGCCAGCTGACAGGACATCTTACAGAAGATTTCACATCTCTGGAAGCTTAGCAA 480
   |||
QY 488 TCTTATTCACACT 501
   |||
Db 481 TCTTATTCACACT 494
   |||

RESULT 10
LOCUS      BG762465
DEFINITION BG762465 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4859042 5',
            mRNA sequence.
ACCESSION  BG762465
VERSION     BG762465.1 GI:14073118
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 905)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1714 row: b column: 03
            High quality sequence stop: 829.
            Location/Qualifiers
            1..905
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
```

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 27 row: D column: 05
High quality sequence stop: 561.
Location/Qualifiers
1. 561

FEATURES
source
1. 561
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="B2N807043-27-D05"
/sex="W"
/lab_host="Top10F"
/clone_lib="B2N807043"
/note="Organ: Brain; Vector: pCNS-D2; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including
EcoRI site by treatment of T4 RNA ligase and the first
strand cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

BASE COUNT 144 a 147 c 123 g 147 t
ORIGIN
Query Match 97.8%; Score 489.8; DB 14; Length 561;
Best Local Similarity 98.6%; Pred. No. 1.5e-131;
Matches 494; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGGACAGACACTCTCGAGGCGCTCTGCG 60
DB 50 CTCCTCAGCTGAACCAATATTACAAATTCGGACAGACACTCTCGAGGCGCTCTGCG 109
QY 61 AGAAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCATG 120
DB 110 AGAAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCATG 169
QY 121 TCTTCTTCTCTTAATTAACCTGGGATGATCTCATTAATAAGAAGATTACTCCCCCTTTA 180
DB 170 TCTTCTTCTCTTAATTAACCTGGGATGATCTCATTAATAAGAAGATTACTCCCCCTTTA 229
QY 181 ACCCAATGTGAGTGGGCGCCCAACGAGCTTACGCGACCTTTGACCCCGAGTTTACCAGAGAGC 240
DB 230 ACCCAATGTGAGTGGGCGCCCAACGAGCTTACGCGACCTTTGACCCCGAGTTTACCAGAGAGC 289
QY 241 CTGTCCCACTCCATTTGGCAAGTCCCTGACAGCGTCTCGTCACAGCCAGCGTCAAGG 300
DB 250 CTGTCCCACTCCATTTGGCAAGTCCCTGACAGCGTCTCGTCACAGCCAGCGTCAAGG 349
QY 301 AAGCTGCGGAGGCTTCTAGGCTTTTCTATGGCCCTCCACGAGCTCTTCTCTGAA 360
DB 350 AAGCTGCGGAGGCTTCTAGGCTTTTCTATGGCCCTCCACGAGCTCTTCTCTGAA 409
QY 361 CCTCTTAGGGCTTGGTTTAAAGATTATGTGTGTTCGGAATGTTTAGTAGGCT 420
DB 410 CCTCTTAGGGCTTGGTTTAAAGATTATGTGTGTTCGGAATGTTTAGTAGGCT 469
QY 421 TTGTGTGAGCGCGCAGCTGACAGGACATCTTACAGAGAAATTGCACATCTCTGGAAGC 480

Db 470 TTTGGTGGAGCGCGCAGCTGACAGACATCTTACAGAGAAATTGGACATCTCTGGAAGC 529
QY 481 TTAGCAATCTTATTGCACACT 501
DB 530 TTAGCAATCTTATTGCACACT 550
RESULT 12
BQ691213 BQ691213 997 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8344598 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6247587
DEFINITION S, mRNA sequence.
ACCESSION BQ691213
VERSION BQ691213.1 GI:21816529
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 997)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csabps-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2386 row: b column: 04
High quality sequence start: 24
High quality sequence stop: 715.
Location/Qualifiers
1. 997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6247587"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_110"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 243 a 255 c 233 g 265 t
ORIGIN
Query Match 97.5%; Score 488.4; DB 13; Length 997;
Best Local Similarity 99.6%; Pred. No. 4.8e-131;
Matches 500; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGGACAGACACTCTCGAGGCGCTCTGCG 60
DB 295 CTCCTCAGCTGAACCAATATTACAAATTCGGACAGACACTCTCGAGGCGCTCTGCG 354
QY 61 AGAAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCATG 120
DB 355 AGAAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCATG 414
QY 121 TCTTCTTCTCTTAATTAACCTGGGATGATCTCATTAATAAGAAGATTACTCCCCCTTTA 180
DB 415 TCTTCTTCTCTTAATTAACCTGGGATGATCTCATTAATAAGAAGATTACTCCCCCTTTA 474
QY 181 ACCCAATGTGAGTGGGCGCCCAACGAGCTTACGCGACCTTTGACCCCGAGTTTACCAGAGAGC 240

```

Db 475 ACCCAATGTGAGTGGGCGCCACAGACCTACGGCAGCTTTGACCCCGAGTTTACCGAAGAGC 534
QY 241 CTGTCCCAACTCCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
Db 535 CTGTCCCAACTCCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 594
QY 301 AAGTCGCCGAGGCTTTCCTAGGCTTTTCCATGCGCCCTCCACGAGCTCTTTCTCTGAA 360
Db 595 AAGTCGCCGAGGCTTTCCTAGGCTTTTCCATGCGCCCTCCACGAGCTCTTTCTCTGAA 654
QY 361 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTATGTTAGTAG-CC 419
Db 655 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTATGTTAGTAGCC 714
QY 420 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAGAGAAATTTGCACATCTCTGGAAG 479
Db 715 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAGAGAAATTTGCACATCTCTGGAAG 774
QY 480 CTAGCAATCTTATTGCACACT 501
Db 775 CTAGCAATCTTATTGCACACT 796

```

RESULT 13

```

AV703365 722 bp mRNA linear EST 09-OCT-2000
LOCUS AV703365 ADB Homo sapiens cDNA clone ADBAEH03 5', mRNA sequence.
ACCESSION AV703365
VERSION AV703365.1 GI:10720693

```

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Peng, Y., Song, H., Huang, Q., Gu, Y., Yang, Y., Gao, G., Xiao

H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,

Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu

G., Hu, R., Chen, J., Chen, Z., and Han, Z.

TITLE

Homo sapiens cDNA ADB clones

JOURNAL

Unpublished

COMMENT

Contact: Zequang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzq@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

Location/Qualifiers

1..722

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="ADBAEH03"

/tissue_type="Adrenal gland"

/dev_stage="Adult"

/lab_host="SOLR"

/clone_lib="ADB"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:

XhoI"

BASE COUNT

174 a 185 c 163 g 199 t 1 others

ORIGIN

Query Match 96.9%; Score 485.6; DB 9; Length 722;

Best Local Similarity 98.0%; Pred. No. 2.8e-130;

Matches 491; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

Db 160 AGAAGACAGSAAAGCGGCTCGGGGCCAAGGATGACTTCATGAGATTAAAGATCATG 219
QY 121 TCTTCTTCTCCCTTAATTAACCTGGATGATCTCATTAATAAGAAGATTACTCCCTCTTAA 180
Db 220 TCTTCTTCTCCCTTAATTAACCTGGATGATCTCATTAATAAGAAGATTACTCCCTCTTAA 279
QY 181 ACCCAATGTGAGTGGGCGCCACAGACCTACGGCAGCTTTGACCCCGAGTTTACCGAAGAGC 240
Db 280 ACCCAATGTGAGTGGGCGCCACAGACCTACGGCAGCTTTGACCCCGAGTTTACCGAAGAGC 319
QY 241 CTGTCCCAACTCCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 300
Db 340 CTGTCCCAACTCCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGG 399
QY 301 AAGTCGCCGAGGCTTTCCTAGGCTTTTCCATGCGCCCTCCACGAGCTCTTTCTCTGAA 360
Db 400 AAGTCGCCGAGGCTTTCCTAGGCTTTTCCATGCGCCCTCCACGAGCTCTTTCTCTGAA 459
QY 361 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTATGTTAGTAGCT 420
Db 460 CCTGTAGGCTTGGTTTAAAGGATTTATGTGTTTCCGAATGTTTATGTTAGTAGCT 519
QY 421 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAGAGAAATTTGCACATCTCTGGAAG 480
Db 520 TTTGTGTGAGCGCCAGCTGACAGGACATCTTACAAGAGAAATTTGCACATCTCTGGAAG 579
QY 481 TTAGCAATCTTATTGCACACT 501
Db 580 TTAGCAATCTTATTGCACACT 600

```

RESULT 14

```

BE730437 912 bp mRNA linear EST 15-SEP-2000
LOCUS BE730437 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3831254 5',
DEFINITION mRNA sequence.
ACCESSION BE730437
VERSION BE730437.1 GI:10144531
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 912)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DYF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (JLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: DCM506 row: 1 column: 15
High quality sequence start: 75
High quality sequence stop: 759.
Location/Qualifiers
1..912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3831254"
/tissue_type="melanotic melanoma"
/lab_host="pH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"

```

```

/note="Organ: Skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in

```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```


```

```
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 227 a 206 c 248 g 231 t

ORIGIN

Query Match 96.3%; Score 482.4; DB 10; Length 912;
Best Local Similarity 99.6%; Pred. No. 2.6e-129;
Matches 494; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCT-GGAGGGCTCTCTG 59
|||||
Db 210 CTCCTCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTGGAGGGCTCTCTG 269
|||||

QY 60 CAGAAGCAGCAGCAACGCGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCAT 119
|||||
Db 270 CAGAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCAT 329
|||||

QY 120 GTCTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAGAAGATTACTCCCGCTTTT 179
|||||
Db 330 GTCTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAGAAGATTACTCCCGCTTTT 389
|||||

QY 180 AACCCAAATGTAGTGGGCCCAAGCAGCTACGCGACTTTGACCCGAGTTTACCGAAGAG 239
|||||
Db 390 AACCCAAATGTAGTGGGCCCAAGCAGCTACGCGACTTTGACCCGAGTTTACCGAAGAG 449
|||||

QY 240 CTTGTCCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 299
|||||
Db 450 CTTGTCCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 509
|||||

QY 300 GAAGCTCCGAGGCTTCTTAGGCTTTCTATAGCGCTCCCGACGAGCTTCTCTCTGA 359
|||||
Db 510 GAAGCTCCGAGGCTTCTTAGGCTTTCTATAGCGCTCCCGACGAGCTTCTCTCTGA 569
|||||

QY 360 ACCCTCTTAGGCGTGGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTAGTATAGCC 419
|||||
Db 570 ACCCTCTTAGGCGTGGTTTAAAGGATTTATGTGTGTTTCCGAATGTTTAGTATAGCC 629
|||||

QY 420 TTTTGTGGAGCGCGCAGCTGACGAGACATCTTTACAAAGAAATTTGCACATCTCTGGAAG 479
|||||
Db 630 TTTTGTGGAGCGCGCAGCTGACGAGACATCTTTACAAAGAAATTTGCACATCTCTGGAAG 689
|||||

QY 480 CTTAGCAATCTTATTG 495
|||||
Db 690 CTTAGCAATCTTATTG 705
|||||

RESULT 15
BQ212589
LOCUS
DEFINITION
AGENCOURT_7675949 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095827
5', mRNA sequence.
BQ212589
VERSION
BQ212589.1 GI:20392956
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 847)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DFP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13368 row: n column: 20

High quality sequence stop: 636.
Location/Qualifiers
1. 847
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6095827"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="organ: skin; Vector: pCMV-SPORT6; Site: 1; Nott;
Site: 2; Salt; Cloned unidirectionally. Primer: oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 194 a 247 c 190 g 216 t

ORIGIN

Query Match 95.9%; Score 480.4; DB 13; Length 847;
Best Local Similarity 98.6%; Pred. No. 9.6e-129;
Matches 495; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTGGAGGGCTCTCTG 60
|||||
Db 265 CTCCTCAGCTGAACCAATATTACAAATTCGCGAAGACACCTCTGGAGGGCTCTCTG 324
|||||

QY 61 AGAAGCAGCAGCAACGCGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCAT 120
|||||
Db 325 AGAAGCAGCAGCAACGCGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAGATCAT 184
|||||

QY 121 TCTCTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAGAAGATTACTCTCTCTGA 180
|||||
Db 385 TCTCTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAGAAGATTACTCTCTCTGA 444
|||||

QY 181 ACCCAAATGTAGTGGGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACGAGAGAG 240
|||||
Db 445 ACCCAAATGTAGTGGGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACGAGAGAG 504
|||||

QY 241 CTGTCCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 300
|||||
Db 505 CTGTCCCCAACTCCATTGGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAG 564
|||||

QY 301 AAGCTCCGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCGACGAGCTCTTCTCTCTGA 360
|||||
Db 565 AAGCTCCGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCGACGAGCTCTTCTCTCTGA 624
|||||

QY 361 CCTGTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTTCGAATGTTTATGTTAGTAC 419
|||||
Db 625 CCTGTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTTCGAATGTTTATGTTAGTAC 684
|||||

QY 420 TTTTGTGGAGCGCGCAGCTGACAGCAGCATCTTTACAAAGAAATTTGCACATCTCTGGAAG 479
|||||
Db 685 TTTTGTGGAGCGCGCAGCTGACAGCAGCATCTTTACAAAGAAATTTGCACATCTCTGGAAG 744
|||||

QY 480 CTTAGCAATCTTATTGACACT 501
|||||
Db 745 CTTAGCAATCTTATTGACACT 766
|||||

Search completed: August 11, 2003, 14:59:33
Job time : 2067 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 13:41:50 ; Search time 73 Seconds
(without alignments)

3029.218 Million cell updates/sec

Title: US-10-000-039A-1_COPY_980_1480

Perfect score: 501

Sequence: 1 CTCCTCAGCTGAACCAAAAT.....TAGCAATCTTATTCACACT 501

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	501	100.0	2370	4	US-09-031-295-1
2	499.4	99.7	2311	2	US-08-712-709-6
3	499.4	98.7	2311	3	US-09-111-444-6
4	499.4	99.7	2311	3	US-09-541-228-6
5	499.4	99.7	2311	4	US-09-016-434-772
6	112	22.4	257	1	US-08-700-575-4
7	48.2	9.6	387	3	US-09-474-922A-2
8	48	9.6	2181	4	US-09-417-197-70
9	48	9.6	2184	4	US-09-417-197-138
10	48	9.6	2610	2	US-09-212-771-1
11	48	9.6	2610	3	US-09-091-058-1
12	46.4	9.3	1599	3	US-09-256-465-1
13	46.4	9.3	1599	4	US-09-167-322-3
14	45	9.0	1273	2	US-08-474-379C-25
15	45	9.0	1273	3	US-09-146-243A-25
16	45	9.0	1273	3	US-08-206-188B-25
17	45	9.0	1276	1	US-07-688-352C-25
18	45	9.0	1276	5	PCT-US91-02714-24
19	44	8.8	1008	4	US-09-394-455-3
20	44	8.8	2146	4	US-09-842-307-1
21	44	8.8	2549	3	US-09-467-082-3
22	44	8.8	2549	4	US-09-394-455-5
23	44	8.8	2599	6	5266464-1
24	44	8.8	2608	4	US-09-394-455-35
25	42.4	8.5	1619	4	US-09-394-455-14
26	41.2	8.2	2196	1	US-08-313-274-1
27	40.8	8.1	1788	4	US-09-417-197-68

28	40.8	8.1	2211	4	US-09-394-455-39
29	39.8	7.9	4041	3	US-09-105-537-36
30	39.8	7.9	36778	3	US-09-105-537-5
31	39.8	7.9	38506	3	US-09-320-878-19
32	39.8	7.9	38506	4	US-09-141-908-1
33	39.8	7.9	38506	4	US-09-657-440-19
34	39.6	7.9	2104	3	US-09-313-930-1
35	39.6	7.9	3255	4	US-09-016-434-1471
36	38.4	7.7	7218	1	US-08-232-463-14
37	38.2	7.6	11220	3	US-09-105-537-32
38	34.2	6.8	9421	2	US-08-370-319C-2
39	34.2	6.8	9421	3	US-09-224-834-2
40	33.8	6.7	2274	4	US-09-772-647-3
C 41	33.6	6.7	19025	4	US-09-849-334-3
C 42	33.4	6.7	1881	4	US-09-434-288-5
43	33	6.6	1636	4	US-09-016-434-1433
44	32.6	6.5	900	4	US-09-393-634-44
45	32.6	6.5	2245	4	US-09-225-749-24

ALIGNMENTS

RESULT 1
US-09-031-295-1
; Sequence 1, Application US/09031295
; Patent No. 6326181
; GENERAL INFORMATION:
; APPLICANT: LANG, Florian
; APPLICANT: WALDEGGER, Tubingen
; TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SOK
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,295
; FILING DATE: 26-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 197-08-173.8
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sandercok, Colin G.
; REGISTRATION NUMBER: 31,298
; REFERENCE/DOCKET NUMBER: 056315/0123
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 43..1335
US-09-031-295-1

Query Match 100.0%; Score 501; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 1.7e-161;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTCCAGCTGAACCAAAATATTACAAATCCGCAAGACACCTCTCTGGAGGCGCTCCTGCG 60
DB CTCTCCAGCTGAACCAAAATATTACAAATCCGCAAGACACCTCTCTGGAGGCGCTCCTGCG 1039
QY 61 AGAAGGACAGCAAAAGCGCTCGGGCCAAAGGATGACTTCATGAGATTAAGAGTCATG 120
DB AGAAGGACAGCAAAAGCGCTCGGGCCAAAGGATGACTTCATGAGATTAAGAGTCATG 1099
QY 121 TCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAATTAATCCCTCTTTTA 180
DB TCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAATTAATCCCTCTTTTA 1159
QY 181 ACCCAAAATGTGAGTGGGCGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACCGAAGG 240
DB ACCCAAAATGTGAGTGGGCGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACCGAAGG 1219
QY 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 300
DB CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 1279
QY 301 AAGCTGCGGAGGCTTCTAGGCTTTCTATGGCGCTCCACAGGACTCTTTCTCTGAA 360
DB AAGCTGCGGAGGCTTCTAGGCTTTCTATGGCGCTCCACAGGACTCTTTCTCTGAA 1339
QY 361 CCTGTTAGGCGCTTGGTTTAAAGGATTTTATGTTGTTCCGATGTTTATAGTAGCCT 420
DB CCTGTTAGGCGCTTGGTTTAAAGGATTTTATGTTGTTCCGATGTTTATAGTAGCCT 1399
QY 421 TTGTTGAGGCGCGGAGCTGACAGGACATCTTACAAGAGAATTTGCACATCTCTGGAAGC 480
DB TTGTTGAGGCGCGGAGCTGACAGGACATCTTACAAGAGAATTTGCACATCTCTGGAAGC 1459
QY 481 TTAGCAATCTTATTGCACACT 501
DB TTAGCAATCTTATTGCACACT 1480

RESULT 2

US-08-712-709-6
; Sequence 6, Application US/0812709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712.709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6

Query Match 99.7%; Score 499.4; DB 2; Length 2311;
Best Local Similarity 99.8%; Pred. No. 5.8e-161;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCTCCAGCTGAACCAAAATATTACAAATCCGCAAGACACCTCTCTGGAGGCGCTCCTGCG 60
DB CTCTCCAGCTGAACCAAAATATTACAAATCCGCAAGACACCTCTCTGGAGGCGCTCCTGCG 1006
QY 61 AGAAGGACAGCAAAAGCGCTCGGGCCAAAGGATGACTTCATGAGATTAAGAGTCATG 120
DB AGAAGGACAGCAAAAGCGCTCGGGCCAAAGGATGACTTCATGAGATTAAGAGTCATG 1066
QY 121 TCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAATTAATCCCTCTTTTA 180
DB TCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAATTAATCCCTCTTTTA 1126
QY 181 ACCCAAAATGTGAGTGGGCGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACCGAAGG 240
DB ACCCAAAATGTGAGTGGGCGCCCAACGAGCTACGCGACTTTTGACCCCGAGTTTACCGAAGG 1186
QY 241 CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 300
DB CTGTCCCAACTCCATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGGCTCAAGG 1246
QY 301 AAGCTGCGGAGGCTTCTAGGCTTTCTATGGCGCTCCACAGGACTCTTTCTCTGAA 360
DB AAGCTGCGGAGGCTTCTAGGCTTTCTATGGCGCTCCACAGGACTCTTTCTCTGAA 1306
QY 361 CCTGTTAGGCGCTTGGTTTAAAGGATTTTATGTTGTTCCGATGTTTATAGTAGCCT 420
DB CCTGTTAGGCGCTTGGTTTAAAGGATTTTATGTTGTTCCGATGTTTATAGTAGCCT 1366
QY 421 TTGTTGAGGCGCGGAGCTGACAGGACATCTTACAAGAGAATTTGCACATCTCTGGAAGC 480
DB TTGTTGAGGCGCGGAGCTGACAGGACATCTTACAAGAGAATTTGCACATCTCTGGAAGC 1426
QY 481 TTAGCAATCTTATTGCACACT 501
DB TTAGCAATCTTATTGCACACT 1447

RESULT 3

US-09-111-444-6
; Sequence 6, Application US/09111444
; Patent No. 6045792
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/111,444
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2311 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-111-444-6

Query Match 99.7%; Score 499.4; DB 3; Length 2311;
Best Local Similarity 99.8%; Pred. No. 5.8e-161;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 60
DB 947 CTCCTCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 1006
QY 61 AGAAGCAGACGACAAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGAGTCATG 120
DB 1007 AGAAGCAGACGACAAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGAGTCATG 1066
QY 121 TCTCTCTCTCTTAATTAACCTGGGATGATCTCAATTAATAAGAGATTAAGAGTCATG 180
DB 1067 TCTCTCTCTCTTAATTAACCTGGGATGATCTCAATTAATAAGAGATTAAGAGTCATG 1126
QY 181 ACCCAATATGAGTGGGCGGCAAGGATGACTTCATGAGGATTAAGAGTCATG 240
DB 1127 ACCCAATATGAGTGGGCGGCAAGGATGACTTCATGAGGATTAAGAGTCATG 1186
QY 241 CTGTCCTCAACTCCATTCGCAAGTCCCTGACAGGCTCCCTGACAGGATTAAGAGTCATG 300
DB 1187 CTGTCCTCAACTCCATTCGCAAGTCCCTGACAGGCTCCCTGACAGGATTAAGAGTCATG 1246
QY 301 AAGCTGCGGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCAAGGATTAAGAGTCATG 360
DB 1247 AAGCTGCGGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCAAGGATTAAGAGTCATG 1306
QY 361 CCTCTTAGGCGTGGTGTAAAGATTTATGTTGTTCCGAATGTTTATGTTAGTGCCT 420
DB 1307 CCTCTTAGGCGTGGTGTAAAGATTTATGTTGTTCCGAATGTTTATGTTAGTGCCT 1366
QY 421 TTGTGTGGGCGGCGGATGACAGGATCTTACAGAGATTTGACATCTCTGGAAGC 480
DB 1367 TTGTGTGGGCGGCGGATGACAGGATCTTACAGAGATTTGACATCTCTGGAAGC 1426
QY 481 TTAGCAATCTTATTCACACT 501
DB 1427 TTAGCAATCTTATTCACACT 1447

RESULT 4
US-09-541-228-6
Sequence 6, Application US/09541228
Patent No. 632077
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/541,228
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/712,709
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0118 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2311 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-541-228-6

Query Match 99.7%; Score 499.4; DB 3; Length 2311;
Best Local Similarity 99.8%; Pred. No. 5.8e-161;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CTCCTCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 60
DB 947 CTCCTCAGCTGAACCAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 1006
QY 61 AGAAGCAGACGACAAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGAGTCATG 120
DB 1007 AGAAGCAGACGACAAAGCGGCTCGGGCCCAAGGATGACTTCATGAGATTAAGAGTCATG 1066
QY 121 TCTCTCTCTCTTAATTAACCTGGGATGATCTCAATTAATAAGAGATTAAGAGTCATG 180
DB 1067 TCTCTCTCTCTTAATTAACCTGGGATGATCTCAATTAATAAGAGATTAAGAGTCATG 1126
QY 181 ACCCAATATGAGTGGGCGGCAAGGATGACTTCATGAGGATTAAGAGTCATG 240
DB 1127 ACCCAATATGAGTGGGCGGCAAGGATGACTTCATGAGGATTAAGAGTCATG 1186
QY 241 CTGTCCTCAACTCCATTCGCAAGTCCCTGACAGGCTCCCTGACAGGATTAAGAGTCATG 300
DB 1187 CTGTCCTCAACTCCATTCGCAAGTCCCTGACAGGCTCCCTGACAGGATTAAGAGTCATG 1246
QY 301 AAGCTGCGGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCAAGGATTAAGAGTCATG 360
DB 1247 AAGCTGCGGAGGCTTCTCTAGGCTTTTCTATGCGCTCCCAAGGATTAAGAGTCATG 1306
QY 361 CCTCTTAGGCGTGGTGTAAAGATTTATGTTGTTCCGAATGTTTATGTTAGTGCCT 420
DB 1307 CCTCTTAGGCGTGGTGTAAAGATTTATGTTGTTCCGAATGTTTATGTTAGTGCCT 1366
QY 421 TTGTGTGGGCGGCGGATGACAGGATCTTACAGAGATTTGACATCTCTGGAAGC 480
DB 1427 TTGTGTGGGCGGCGGATGACAGGATCTTACAGAGATTTGACATCTCTGGAAGC 1426

Db 1367 TTGTGGAGCCGCGAGCTGACAGACATCTTACAAGAGATTTCACATCTCTGGAAGC 1426
QY 481 TTAGCAATCTTATTGCACACT 501
Db 1427 TTAGCAATCTTATTGCACACT 1447

RESULT 5

US-09-016-434-772
; Sequence 772, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
US-09-016-434-772

Query Match 99.7%; Score 499.4; DB 4; Length 2311;
Best Local Similarity 99.8%; Pred. No. 5.8e-161;
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CTCCTCAGCTGAACAAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 60
Db 947 CTCCTCAGCTGAACAAATATTACAAATTCGCAAGACACCTCTCGAGGCGCTCTGCG 1006
QY 61 AGAAGACAGGCAAGCGCTCGGGCCCAAGGATGACTTCATGAGGATTAAGATCATG 120
Db 1007 AGAAGACAGGCAAGCGCTCGGGCCCAAGGATGACTTCATGAGGATTAAGATCATG 1066
QY 121 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 180
Db 1067 TCTTCTCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1126
QY 181 ACCCAATGTGAGTGGGCGCCACAGCTAGCGACTTTTGACCCCGAGTTTACCGAAGAGC 240
|||||

Db 1127 ACCCAAAATGTGAGTGGGCGCCCAAGCACCTACGCGACTTTTGACCCCGAGTTTATCCGAAGAGC 1186
QY 241 CTGTCCCAACTCCATTGSCAAGTCCCTCGTACAGAGCTCTCGTACAGCCAGCGTCAAGG 300
|||||
Db 1187 CTGTCCCAACTCCATTGSCAAGTCCCTCGTACAGAGCTCTCGTACAGCCAGCGTCAAGG 1246
QY 301 AAGCTGCGGAGGCTTTCCTAGGCTTTTCTATGCGCTCCCAAGGACTCTTTTCCGTCGAA 360
|||||
Db 1247 AAGCTGCGGAGGCTTTCCTAGGCTTTTCTATGCGCTCCCAAGGACTCTTTTCCGTCGAA 1306
QY 361 CCGTGTAGGCGCTGCTTTTAAAGGATTTTATGTGTGTTCCGAATGTTTATGTTAGCT 420
|||||
Db 1307 CCGTGTAGGCGCTGCTTTTAAAGGATTTTATGTGTGTTCCGAATGTTTATGTTAGCT 1366
QY 421 TTTGTGAGCGCGCGAGCTGACAGACATCTTACAAGAGATTTCACATCTCTGGAAGC 480
|||||
Db 1367 TTTGTGAGCGCGCGAGCTGACAGACATCTTACAAGAGATTTCACATCTCTGGAAGC 1426
QY 481 TTAGCAATCTTATTGCACACT 501
|||||
Db 1427 TTAGCAATCTTATTGCACACT 1447

RESULT 6

US-08-700-575-4
; Sequence 4, Application US/08700575
; Patent No. 5817479
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Philip R.
; APPLICANT: Wilde, Craig G.
; TITLE OF INVENTION: NOVEL HUMAN KINASE HOMOLOGS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,575
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J
; REGISTRATION NUMBER: 36749
; REFERENCE/DOCKET NUMBER: SP-100 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Thp-1 Phorbol I/PS
; CLONE: 12702
US-08-700-575-4

Query Match 22.4%; Score 112; DB 1; Length 257;
Best Local Similarity 93.1%; Pred. No. 1.3e-28;
Matches 162; Conservative 0; Mismatches 5; Indels 7; Gaps 4;

;; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-1 EXPRESSION
;; FILE REFERENCE: RTS-0034
;; CURRENT APPLICATION NUMBER: US/09/212,771
;; CURRENT FILING DATE: 1998-12-16
;; NUMBER OF SEQ ID NOS: 47
;; SEQ ID NO 1
;; LENGTH: 2610
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (199)..(1641)
US-09-212-771-1

Query Match 9.6%; Score 48; DB 3; Length 2610;
Best Local Similarity 50.9%; Pred. No. 5.5e-06;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 54 CTCCTGCAGAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAG 113
Db 1348 CTCAGAAGGACCCCAAGCAGAGGCTTGGCGGGCTCCGAGGACGCCAAGGAGATCATG 1407
QY 114 AGTCATGCTCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAAGATTACTCCC 173
Db 1408 CAGCATGCTCTTCTTGGCGGTATCGTGTGGCAGCAGTGTACGAGAAAGCTCAGCCCA 1467
QY 174 CCTTTTAACCCAAATGTGAGTGGGCCCCAACGAGCTACGCACTTTGACCCCGAGTTTACC 233
Db 1468 CCTTCAAGCCCCAGGTCAAGTCCGAGACTGACACCCAGGTATTTTGTGAGGAGTTACG 1527
QY 234 GAAGAGCTGTCCTCCCACTCCATTGGCAAGTCCCTGCACAGCGT 277
Db 1528 GCCCAGATGATCACCATCACACCCACTGCACCAAGATGACAGCAT 1571

RESULT 11
US-09-091-058-1
;; Sequence 1, Application US/09091058
;; Patent No. 6054285
;; GENERAL INFORMATION:
;; APPLICANT: Hemmings, Brian A.
;; APPLICANT: Frech, Matthias
;; TITLE OF INVENTION: Screening Method
;; FILE REFERENCE: 4-20683/A/20684/PCT
;; CURRENT APPLICATION NUMBER: US/09/091,058
;; CURRENT FILING DATE: 1998-06-10
;; EARLIER APPLICATION NUMBER: PCT/EP96/04814
;; EARLIER FILING DATE: 1996-11-05
;; EARLIER APPLICATION NUMBER: 9525703.6
;; EARLIER FILING DATE: 1995-12-15
;; NUMBER OF SEQ ID NOS: 23
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 2610
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (199)..(1641)
US-09-091-058-1

Query Match 9.6%; Score 48; DB 3; Length 2610;
Best Local Similarity 50.9%; Pred. No. 5.5e-06;
Matches 114; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 54 CTCCTGCAGAGGACAGACAAAGCGGCTCGGGGCCAAGGATGACTTCATGGAGATTAAAG 113
Db 1348 CTCAGAAGGACCCCAAGCAGAGGCTTGGCGGGCTCCGAGGACGCCAAGGAGATCATG 1407
QY 114 AGTCATGCTCTTCTCTCTTAATTAACCTGGGATGATCTCATTAAAGAAAGATTACTCCC 173
Db 1408 CAGCATGCTCTTCTTGGCGGTATCGTGTGGCAGCAGCTGTACGAGAAAGCTCAGCCCA 1467

QY 174 CCTTTTAACCCAAATGTGAGTGGGCCCCAACGAGCTACGCACTTTGACCCCGAGTTTACC 233
Db 1468 CCTTCAAGCCCCAGGTCAAGTCCGAGACTGACACCCAGGTATTTTGTGAGGAGTTACG 1527
QY 234 GAAGAGCTGTCCTCCCACTCCATTGGCAAGTCCCTGCACAGCGT 277
Db 1528 GCCCAGATGATCACCATCACACCCACTGCACCAAGATGACAGCAT 1571
RESULT 12
US-09-256-465-1
;; Sequence 1, Application US/09256465
;; Patent No. 6043090
;; GENERAL INFORMATION:
;; APPLICANT: Brett P. Monia
;; APPLICANT: Lex M. Cowser
;; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-2 EXPRESSION
;; FILE REFERENCE: RTS-0035
;; CURRENT APPLICATION NUMBER: US/09/256,465
;; CURRENT FILING DATE: 1999-02-23
;; NUMBER OF SEQ ID NOS: 47
;; SEQ ID NO 1
;; LENGTH: 1599
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (88)..(1533)
US-09-256-465-1

Query Match 9.3%; Score 46.4; DB 3; Length 1599;
Best Local Similarity 52.5%; Pred. No. 1.4e-05;
Matches 126; Conservative 0; Mismatches 111; Indels 3; Gaps 1;
QY 41 CCTCTGGAGGGCTCTCTGCGAGAGGACAGACAAAGCGGT---CGGGCCCAAGGATGA 97
Db 1224 CTGCTTGGGCTGCTTAAGAAGGACCCCAAGCAGAGGCTTGGTGGGGGGCCGACGCA 1283
QY 98 CTTCATGGAGATTAAAGATCATGCTCTTCTCTCTTAATTAACCTGGGATGATCTCATTA 157
Db 1284 TGCCAGGAGGTGATGGAGCAGAGTCTTCTCAGCATCACTGGCAGGACGTGCTCCA 1343
QY 158 TAAGAAGATTACTCCCTTTTAACCCAAATGTGAGTGGGGCCCAAGGCTTACGCACTT 217
Db 1344 GAAGAAGCTCTCGCAGCCCTTCAAACCTCAGCTCAGCTCAGCTCGAGGTCGACACAGG 1403
QY 218 TGACCCCGAGTTTACCGAAGAGCCTGTCGCCCAACTCCATTGGCAAGTCCCTGCACG 277
Db 1404 CGATGATGAATTTACCGCCAGTCCATCAATCAACACCCCTGACCGCTATGACAGCT 1463

RESULT 13
US-09-167-322-3
;; Sequence 3, Application US/09167322
;; Patent No. 6365151
;; GENERAL INFORMATION:
;; APPLICANT: Allegheny University of the Health
;; Sciences, Halpern, Michael S.
;; England, James M.
;; TITLE OF INVENTION: CANCER VACCINE
;; NUMBER OF SEQUENCES: 14
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.
;; STREET: Suite 1800, Two Penn Center Plaza
;; CITY: Philadelphia
;; STATE: PA
;; COUNTRY: USA
;; ZIP: 19102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30


```
;
; REGISTRATION NUMBER: 36,107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1273 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..504
; US-09-146-249A-25

Query Match          9.0%; Score 45; DB 3; Length 1273;
Best Local Similarity 57.1%; Pred. No. 3.7e-05;
Matches 101; Conservative 0; Mismatches 75; Indels 1; Gaps 1;

QY 93 GATGACTTCATGGAGATTAAAGATCAAGTCTCTCTCTTAATTAACCTGGGATGATCTC 152
    |||| | || |||| || || || || || || || || || || || || || ||
Db 922 GATGGAGTTGAAGAAATTAAGAGACATTCATTTTCTCAACGATGACTGGAATAACTG 981
    || || || || || || || || || || || || || || || || || || || ||
QY 153 ATTAATAAGAGATTACTCCCTTTTAAACCCAAATGTGAGTGGGCCCAACGAGCTACGG 212
    || || || || || || || || || || || || || || || || || || || ||
Db 982 TATAGAGAGAA-ATTCATCCGCCATTTAACCTGCAACGGCGAGCCTGAAGTACATTC 1040
    || || || || || || || || || || || || || || || || || || || ||
QY 213 CACTTTGACCCCGAGTTTACCGAAGAGCCTGTCCCCCAACTCCATTGGCAAGTCCCT 269
    || || || || || || || || || || || || || || || || || || || ||
Db 1041 TATTTGATCCTGAGTTTACTGCAAAACCTCCCAAGATTACCTGGCATTCACCT 1097
    || || || || || || || || || || || || || || || || || || || ||
```

Search completed: August 11, 2003, 15:00:58
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 01:27:02 : Search time 8486 seconds
(without alignments)
11425.385 Million cell updates/sec

Title: US-10-000-039a-1

Perfect score: 2370

Sequence: 1 CACGAGGAGCGCTACGTC.....AAAAAAAAAAAAAAAAAAAA 2370

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hug.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vi.*

29: em_hug_hum.*

30: em_hug_inv.*

31: em_hug_other.*

32: em_hug_mus.*

33: em_hug_pln.*

34: em_hug_rod.*

35: em_hug_mam.*

36: em_hug_vrt.*

37: em_sy.*

38: em_hug_hum.*

39: em_hug_mus.*

40: em_hug_pln.*

41: em_hug_rod.*

42: em_hug_mam.*

43: em_hug_vrt.*

44: em_sy.*

45: em_hug_hum.*

46: em_hug_mus.*

47: em_hug_pln.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	2370	100.0	2370	6	AR179441	AR179441 Sequence
2	2370	100.0	2370	6	AX002570	AX002570 Sequence
3	2370	100.0	2370	6	AX337834	AX337834 Sequence
4	2370	100.0	2370	6	AX411211	AX411211 Sequence
5	2370	100.0	2370	9	HSRNASPEK	Y10032 H. sapiens m
6	2328.2	98.2	2382	9	AF153609	AF153609 Homo sapi
7	2316.2	97.7	2382	9	BC001263	BC001263 Homo sapi
8	2312.6	97.6	2354	6	AX553543	AX553543 Sequence
9	2311.6	97.5	2364	9	AK096509	AK096509 Homo sapi
10	2262.8	95.5	2311	6	AR151390	AR151390 Sequence
11	2262.8	95.5	2311	6	AR270209	AR270209 Sequence
12	2261.2	95.4	2311	6	BD021928	BD021928 Human pro
13	2201	92.9	3196	6	AX713538	AX713538 Sequence
14	2201	92.9	3196	9	AK055077	AK055077 Homo sapi
15	2140.8	90.3	2281	6	AX017284	AX017284 Sequence
16	2140.8	90.3	2281	6	AX524988	AX524988 Sequence
17	2140.8	90.3	2281	6	BD134455	BD134455 Human tuc
18	1539.2	64.9	2423	10	BC005720	BC005720 Mus muscu
19	1526.4	64.4	2429	10	AF139638	AF139638 Mus muscu
20	1522.4	64.2	2426	10	AF205855	AF205855 Mus muscu
21	1492.4	63.0	2432	6	AX056375	AX056375 Sequence
22	1473.6	62.2	1920	4	AF139639	AF139639 Oryctolap
23	1419.6	59.9	2435	10	RATSGPK	L01624 Rattus norv
24	1296	54.7	1296	6	AX056374	AX056374 Sequence
25	1179.4	49.8	5718	9	HSSGK	AJ000512 Homo sapi
26	1152.6	48.6	113673	9	AL135839	AL135839 Human DNA
27	911.2	38.4	1659	10	BC002222	BC002222 Mus muscu
28	852	35.9	1417	5	AF057138	AF057138 Xenopus l
29	786	33.2	216056	2	AC103256	AC103256 Rattus no
30	779.4	32.9	2339	5	BC052134	BC052134 Danio rer
31	730.4	30.8	2470	5	SAC223715	AJ223715 Squalus a
32	705	29.7	3105	5	SAC223716	AJ223716 Squalus a
33	613.8	25.9	183140	10	AC124486	AC124486 Mus muscu
34	587.8	24.8	208405	2	AC114405	AC114405 Mus muscu
35	534.8	22.6	1803	10	AF312007	AF312007 Mus muscu
36	534.8	22.6	2494	10	BC018363	BC018363 Mus muscu
37	527.2	22.2	1812	6	AX056472	AX056472 Sequence
38	527.2	22.2	1812	6	AX399684	AX399684 Sequence
39	527.2	22.2	1812	9	AF169034	AF169034 Homo sapi
40	527.2	22.2	1815	9	BC014037	BC014037 Homo sapi
41	527.2	22.2	1834	6	AX026166	AX026166 Sequence
42	527.2	22.2	2146	6	AX026164	AX026164 Sequence
43	527.2	22.2	2146	9	AF186470	AF186470 Homo sapi
44	522	22.0	1768	9	BC015326	BC015326 Homo sapi
45	522	22.0	2391	9	AF169035	AF169035 Homo sapi

ALIGNMENTS

RESULT 1
AR179441
LOCUS AR179441
DEFINITION Sequence 1 from patent US 6326181.
ACCESSION AR179441
VERSION AR179441.1
KEYWORDS GI:20220996
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2370)
AUTHORS Lang, F. and Waldegger, S.
TITLE Cell volume-regulated human kinase h-sgk
JOURNAL Patent: US 6326181-A | 04-DEC-2001;
FEATURES Location/Qualifiers

linear PAT 20-APR-2002

[illegible]

|||||
2041 ATGCCCCAGTTGTGTCAGTCAGAGCGCTGGTGTCTTTTCATCTCTTTAAAAATGTCACCTGTAA 2100
|||||
2101 AATGGGCAATATATATGTTTTTTTTTTTTCATCTCTGATATATGTTATGTAAG 2160
|||||
2101 AATGGGCAATATATGTTTTTTTTTTTTCATCTCTGATATATGTTATGTAAG 2160
|||||
2161 AACGCTGTACATTTGGGTTATAACACTAGTATATTTAAACTTACAGGCTTATTTGTAATG 2220
|||||
2161 AACGCTGTACATTTGGGTTATAACACTAGTATATTTAAACTTACAGGCTTATTTGTAATG 2220
|||||
2221 TAAACCCACATTTTAACTGATCTGTAATTAACATGGTTATATATAGTACATCTTCCCTC 2280
|||||
2221 TAAACCCACATTTTAACTGATCTGTAATTAACATGGTTATATATAGTACATCTTCCCTC 2280
|||||
2281 ATCCCATCACACACTTTTTTTTGTGTGTGATAAAGTATTTGTTTGCATATAAACCCTT 2340
|||||
2281 ATCCCATCACACACTTTTTTTTGTGTGTGATAAAGTATTTGTTTGCATATAAACCCTT 2340
|||||
2341 GAAAAATAAAAAAATAAAAAA 2370
2341 GAAAAATAAAAAAATAAAAAA 2370

RESULT 2
AX002570 2370 bp DNA Linear PAT 10-MAR-2000
LOCUS
DEFINITION Sequence 1 from Patent EP0861896.
ACCESSION AX002570
VERSION AX002570.1 GI:7242111
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lang,F.P. and Waldegger,S.D.
TITLE Cell volume regulated human kinase h-sgk
JOURNAL Patent: EP 0861896-A 1 02-SEP-1998;
DADE BEHRING MARBURG GMBH (DE)
FEATURES
Location/Qualifiers
1..2370
/organism="Homo sapiens"
/mol_type="genomic DNA"
/strain="CAUCASIAN"
/db_xref="taxon:9606"
/clone="HSRNAS7PK"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN

Query Match 100.0%; Score 2370; DB 6; Length 2370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2370; Conserved 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACGAGGAGGCGCTAAACGCTCTTCTGCTCCCGGGTGGTGATGACGGTCAAAACTGAG 60
DB 1 CACGAGGAGGCGCTAAACGCTCTTCTGCTCCCGGGTGGTGATGACGGTCAAAACTGAG 60
QY 61 GCTGCTAAGGCAACCCCTCACTTACCCAGGATGAGGGGATGGTGGAATTTCTCATCGCT 120
DB 61 GCTGCTAAGGCAACCCCTCACTTACCCAGGATGAGGGGATGGTGGAATTTCTCATCGCT 120
QY 121 TTCATGAAGCAGAGGAGGATGGTCTGAACGCTTTATTCAGAGATTGCCAATTAACCTCC 180
DB 121 TTCATGAAGCAGAGGAGGATGGTCTGAACGCTTTATTCAGAGATTGCCAATTAACCTCC 180
QY 181 TATGATGCAACACCCCTGAAGTTTCAGTCCCATCTTTGAAGATCTCCCAACCTCAGGAGCT 240
DB 181 TATGATGCAACACCCCTGAAGTTTCAGTCCCATCTTTGAAGATCTCCCAACCTCAGGAGCT 240
QY 241 GAGCTTATGATGCAACCCCTTCTCTCCACCAAGTCTTCTTCAGCAATCAACCTTGGC 300
DB 241 GAGCTTATGATGCAACCCCTTCTCTCCACCAAGTCTTCTTCAGCAATCAACCTTGGC 300

QY 301 CCCTGCTCCAACTCCTCATGCTTAAACCATCTGACTTTTCTTCTTGAAGTGATCGGAAG 360
DB 301 CCCTGCTCCAACTCCTCATGCTTAAACCATCTGACTTTTCTTCTTGAAGTGATCGGAAG 360
QY 361 GGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAGGAGAGAGTGTCTTATGACATC 420
DB 361 GGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAGGAGAGAGTGTCTTATGACATC 420
QY 421 AAAGCTTTTACAGAGAGAGCAATCTGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 AAAGCTTTTACAGAGAGAGCAATCTGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 CGSAAATGTTCTGTTGAAGAATGTGAAGCACCCCTTTCTCTGTTGGGCTTCTTCTTC 540
DB 481 CGSAAATGTTCTGTTGAAGAATGTGAAGCACCCCTTTCTCTGTTGGGCTTCTTCTTC 540
QY 541 CAGACTGCTGACAAATTTGACTTTTCTAGACTACATTAATGTTGAGAGAGTTCTTCTAC 600
DB 541 CAGACTGCTGACAAATTTGACTTTTCTAGACTACATTAATGTTGAGAGAGTTCTTCTAC 600
QY 601 CATCTCCAGAGGGAACGCTGCTTCTGGAACACGGGCTGCTTTCTATGCTGCTGAATA 660
DB 601 CATCTCCAGAGGGAACGCTGCTTCTGGAACACGGGCTGCTTTCTATGCTGCTGAATA 660
QY 661 GCCAGTGGCTTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GCCAGTGGCTTGGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AATATTTTGTAGATTTCACAGGAGACATTTGCTTACTGATTTGCGAGCTGCTGAAGAG 780
DB 721 AATATTTTGTAGATTTCACAGGAGACATTTGCTTACTGATTTGCGAGCTGCTGAAGAG 780
QY 781 AACATTGAACACACAGCAGCAACATCCACCTTCTGTCGACGCGGATATCTGCACT 840
DB 781 AACATTGAACACACAGCAGCAACATCCACCTTCTGTCGACGCGGATATCTGCACT 840
QY 841 GAGTGTCTTATAGACAGCCTTATTCACAGGACTGTGGACTGTGGTGGTGGTGGTGGTGGT 900
DB 841 GAGTGTCTTATAGACAGCCTTATTCACAGGACTGTGGACTGTGGTGGTGGTGGTGGTGGT 900
QY 901 TTGTATGAGATGCTGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 901 TTGTATGAGATGCTGTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 961 GACAACTTCTGAACAGCCTTCTCCAGCTGAAACCAATATTAACAAATTCGCAAGACAC 1020
DB 961 GACAACTTCTGAACAGCCTTCTCCAGCTGAAACCAATATTAACAAATTCGCAAGACAC 1020
QY 1021 CTCTGGAGGCGCTCTCCAGAGACAGGACAGGAGGCTGCGGGGCGCAAGAGTGAATGATC 1080
DB 1021 CTCTGGAGGCGCTCTCCAGAGAGGACAGGAGGCTGCGGGGCGCAAGAGTGAATGATC 1080
QY 1081 ATGAGATTAAGAGTCACTGCTTCTCTCTTAACTGAGTGAATGATCTCATTAAATAG 1140
DB 1081 ATGAGATTAAGAGTCACTGCTTCTCTCTTAACTGAGTGAATGATCTCATTAAATAG 1140
QY 1141 AAGATTACTCCCCCTTTTAAACCAATGAGTGGGCGGCAACAGAGTACGGACTTTGAC 1200
DB 1141 AAGATTACTCCCCCTTTTAAACCAATGAGTGGGCGGCAACAGAGTACGGACTTTGAC 1200
QY 1201 CCGGAGTTTACCGAGAGGCTGCTCCCACTCCATTTGGCAAGTCCCTGACAGGCTGCTC 1260
DB 1201 CCGGAGTTTACCGAGAGGCTGCTCCCACTCCATTTGGCAAGTCCCTGACAGGCTGCTC 1260
QY 1261 GTCACAGCAGCTCAAGGAGGCTTCTGAGGCTTTCTAGGCTTTTCTATGCTGCTGCTC 1320
DB 1261 GTCACAGCAGCTCAAGGAGGCTTCTGAGGCTTTCTAGGCTTTTCTATGCTGCTGCTC 1320
QY 1321 ACGGACTCTTTCCTCTGAACCCCTGTTAGGGCTTGGTTTTAAAGGATTTTATGCTGCTT 1380
DB 1321 ACGGACTCTTTCCTCTGAACCCCTGTTAGGGCTTGGTTTTAAAGGATTTTATGCTGCTT 1380

QY 1381 CGAATGTTTATAGTCTTTTGGTGGAGCGCCAGCTGACAGACATCTTACAAGAA 1440
 DB 1381 CGAATGTTTATAGTCTTTTGGTGGAGCGCCAGCTGACAGACATCTTACAAGAA 1440
 QY 1441 TTGACACATCTCTGGAAGCTTAGCAATCTTATTCACACTGTTCCGCTGGAAATTTTGA 1500
 DB 1441 TTGACACATCTCTGGAAGCTTAGCAATCTTATTCACACTGTTCCGCTGGAAATTTTGA 1500
 QY 1501 GAGCACATCTCTGAGCTAGCTCATGAGGTTTCATCTTTTCTTCTTCCCTCCAGCTGG 1560
 DB 1501 GAGCACATCTCTGAGCTAGCTCATGAGGTTTCATCTTTTCTTCTTCCCTCCAGCTGG 1560
 QY 1561 TCTATCTCTGAACAGCGCTTAGAGTCCCGCTTAGACGAGGAGGAGCTTTCTGTAGA 1620
 DB 1561 TCTATCTCTGAACAGCGCTTAGAGTCCCGCTTAGACGAGGAGGAGCTTTCTGTAGA 1620
 QY 1621 AAGCGAGCTGTTCTAAAAAGGCTCTCTGCAGATCTGCTGGGCTGTGATGACGATAT 1680
 DB 1621 AAGCGAGCTGTTCTAAAAAGGCTCTCTGCAGATCTGCTGGGCTGTGATGACGATAT 1680
 QY 1681 TATGAATCTGCTCTGAGAGATGTTGTTAGCTCCAAAGCTTTTCTCTATCGCAGTG 1740
 DB 1681 TATGAATCTGCTCTGAGAGATGTTGTTAGCTCCAAAGCTTTTCTCTATCGCAGTG 1740
 QY 1741 TTTGAGTCTTTATTTTCCCTTGGGATATCTGCTGGAACCGCTGCTGAGTGTGAT 1800
 DB 1741 TTTGAGTCTTTATTTTCCCTTGGGATATCTGCTGGAACCGCTGCTGAGTGTGAT 1800
 QY 1801 GCTGATCACAGATGATTTGTTTATAAGCATCAATGCTGACACTTGCAGGACACTACAAC 1860
 DB 1801 GCTGATCACAGATGATTTGTTTATAAGCATCAATGCTGACACTTGCAGGACACTACAAC 1860
 QY 1861 GTGGGACATGTTGTTCTCCATATTTGGAAGTAAATTTATGCTGACACTTTTGT 1920
 DB 1861 GTGGGACATGTTGTTCTCCATATTTGGAAGTAAATTTATGCTGACACTTTTGT 1920
 QY 1921 AAGATACGCTTAATTAACATAAATTTATGAAATGCTTGCATGACTGCTATTCAGATG 1980
 DB 1921 AAGATACGCTTAATTAACATAAATTTATGAAATGCTTGCATGACTGCTATTCAGATG 1980
 QY 1981 CCTAAGAAGCATGCTGCTACAAATTTCTTATTTTGAAGGCTTTTATGGACCA 2040
 DB 1981 CCTAAGAAGCATGCTGCTACAAATTTCTTATTTTGAAGGCTTTTATGGACCA 2040
 QY 2041 ATGCCCCAGTCTGAGTCAGAGCCGCTGCTGTTTTCATGTTTAAATGTCACCTGTAA 2100
 DB 2041 ATGCCCCAGTCTGAGTCAGAGCCGCTGCTGTTTTCATGTTTAAATGTCACCTGTAA 2100
 QY 2101 ATGGGCAATATTTATGTTTATTTTGGTCTGATATTTGATGTTATGATTAAG 2160
 DB 2101 ATGGGCAATATTTATGTTTATTTTGGTCTGATATTTGATGTTATGATTAAG 2160
 QY 2161 AAGCTCTGACATTTGTTTATTAACACTAGTATATTAACACTTACAGGCTTATTTGTAATG 2220
 DB 2161 AAGCTCTGACATTTGTTTATTAACACTAGTATATTAACACTTACAGGCTTATTTGTAATG 2220
 QY 2221 TAAACACCATTTTATGTAATGTAATTAACACTGTTATTAACACTTACAGGCTTATTTGTAATG 2280
 DB 2221 TAAACACCATTTTATGTAATGTAATTAACACTGTTATTAACACTTACAGGCTTATTTGTAATG 2280
 QY 2281 ATCCCATCACACACTTTTGTGTGATTAACACTGTTATTAACACTTACAGGCTTATTTGTAATG 2340
 DB 2281 ATCCCATCACACACTTTTGTGTGATTAACACTGTTATTAACACTTACAGGCTTATTTGTAATG 2340
 QY 2341 GAAAAATAAAAAATAAAAAAATAAAAAA 2370
 DB 2341 GAAAAATAAAAAATAAAAAAATAAAAAA 2370

RESULT 3
 AX337834 2370 bp DNA linear PAT 09-JAN-2002
 LOCUS
 DEFINITION Sequence 8343 from Patent WO0194629.

ACCESSION AX337834
 VERSION AX337834.1 GI:18128553
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrikan, S., Soppet, D.R. and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 8343 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 Location/Qualifiers
 1..2370
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 636 a 517 c 513 g 704 t
 ORIGIN
 Query Match 100.0%; Score 2370; DB 6; Length 2370;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CACGAGGAGCGCTAACGCTCTTCTGCTCCCGCGGTGGTGTGATGACGGTCAAACTGAG 60
 DB 1 CACGAGGAGCGCTAACGCTCTTCTGCTCCCGCGGTGGTGTGATGACGGTCAAACTGAG 60
 QY 61 GCTGTATAGGCGCACCTCTACCTTACTCCAGGATGAGGGGATGGTGGCAATTTCTATGCT 120
 DB 61 GCTGTATAGGCGCACCTCTACCTTACTCCAGGATGAGGGGATGGTGGCAATTTCTATGCT 120
 QY 121 TTCATGACGAGGAGGATGGTGTGAACGACTTTTATTCAGAGATTTGCCAATACTATC 180
 DB 121 TTCATGACGAGGAGGATGGTGTGAACGACTTTTATTCAGAGATTTGCCAATACTATC 180
 QY 181 TATGATGCAAAACCCCTGAAGCTTCAGTCCCATCTTGAAGATCTCCCAACCTCAGAGCT 240
 DB 181 TATGATGCAAAACCCCTGAAGCTTCAGTCCCATCTTGAAGATCTCCCAACCTCAGAGCT 240
 QY 301 CCGTCTGCAATCTCTCATGCTAAACCATCTGACTTTCTACTTTCTTGAAGTCAATGGAAG 360
 DB 301 CCGTCTGCAATCTCTCATGCTAAACCATCTGACTTTCTACTTTCTTGAAGTCAATGGAAG 360
 QY 361 GGCAGTTTGGAAAGGCTTCTTACGACAGACACAGCCAGAGAGCTTCTATGCAATC 420
 DB 361 GGCAGTTTGGAAAGGCTTCTTACGACAGACACAGCCAGAGAGCTTCTATGCAATC 420
 QY 421 AAGATTTTACAGAGAAAGCAATCTGAAAGAAAGAGAGAGAGATATATGTCGAG 480
 DB 421 AAGATTTTACAGAGAAAGCAATCTGAAAGAAAGAGAGAGAGATATATGTCGAG 480
 QY 481 CGGAATGTTCTGTTGAAGAAATGTGAACACCCCTTCTGTTGGGCTTCTACTTCTTTC 540
 DB 481 CGGAATGTTCTGTTGAAGAAATGTGAACACCCCTTCTGTTGGGCTTCTACTTCTTTC 540
 QY 541 CAGAGTGTGCAAAATCTGCTTCTCTAGACTACATTAATGCTGAGAGTGTCTCTAC 600
 DB 541 CAGAGTGTGCAAAATCTGCTTCTCTAGACTACATTAATGCTGAGAGTGTCTCTAC 600
 QY 601 CATCTCCAGAGGAGCGCTGCTTCTGGAACACCGGCTCGTTTCTATGCTGTGAATA 660
 DB 601 CATCTCCAGAGGAGCGCTGCTTCTGGAACACCGGCTCGTTTCTATGCTGTGAATA 660
 QY 661 GCCAGTGTGCTGGCTACCTGCTTCTGCAACATCTGCTTATAGAGACTTAAACACAG 720
 DB 661 GCCAGTGTGCTGGCTACCTGCTTCTGCAACATCTGCTTATAGAGACTTAAACACAG 720

	QY	1801	GCTGTATCAGATGGAATTGGTTTATAAAGCATCAATGTGCACACTTCGACAGCACATTAAC	1860
	DB	1801	GCTGTATCAGATGGAATTGGTTTATAAAGCATCAATGTGCACACTTCGACAGCACATTAAC	1860
	QY	1861	GTGGGACATTTGTTTGTTCTCCCATATTTTGGAAAGATAAATTTATSTGTAGACATTTTGT	1920
	DB	1861	GTGGGACATTTGTTTGTTCTCCCATATTTTGGAAAGATAAATTTATSTGTAGACATTTTGT	1920
	QY	1921	AAGATACGGTTTAACTAAAATTTATGAAATGGTCTTGCATGACTCGTATTCAGATG	1980
	DB	1921	AAGATACGGTTTAACTAAAATTTATGAAATGGTCTTGCATGACTCGTATTCAGATG	1980
	QY	1981	CCTAAAGAAAGCATTTGCTGCTACAAATATTTCTATTTTTTAGAAAGGTTTATGAGACA	2040
	DB	1981	CCTAAAGAAAGCATTTGCTGCTACAAATATTTCTATTTTTTAGAAAGGTTTATGAGACA	2040
	QY	2041	ATGCCCCAGTTGTCAGTCAGAGCGCTTGTGGTGTGTTTTCATTTTAAATGTACCCTGTAA	2100
	DB	2041	ATGCCCCAGTTGTCAGTCAGAGCGCTTGTGGTGTGTTTTCATTTTAAATGTACCCTGTAA	2100
	QY	2101	AATGGGCATTAATTTATGTTTTTTTTPTTGCATTCCTGATAATTTGATTTGATAAAG	2160
	DB	2101	AATGGGCATTAATTTATGTTTTTTTTPTTGCATTCCTGATAATTTGATTTGATAAAG	2160
	QY	2161	AACGCTCTACATTTGGGTATTAACACTAGTATATTTAAACTTACAGCTTATTTGTAATG	2220
	DB	2161	AACGCTCTACATTTGGGTATTAACACTAGTATATTTAAACTTACAGCTTATTTGTAATG	2220
	QY	2221	TAAACCCACATTTTAATGTACTGTAATTAACATGGTTTATAATACGTAACATCTTCTTCT	2280
	DB	2221	TAAACCCACATTTTAATGTACTGTAATTAACATGGTTTATAATACGTAACATCTTCTTCT	2280
	QY	2281	ATCCCATCACACAACATTTTGTGTGTGATAACTGATTTTGGTTTGGCAATAAAACTT	2340
	DB	2281	ATCCCATCACACAACATTTTGTGTGTGATAACTGATTTTGGTTTGGCAATAAAACTT	2340
	QY	2341	CAAAATAAAAAAAAAAAAAAAAAAAAAA 2370	
	DB	2341	CAAAATAAAAAAAAAAAAAAAAAAAAAA 2370	
	RESULT 4			
	AX411211			
	LOCUS	AX411211	Sequence 3858 from Patent WO0229103.	DNA 1linear PAT 14-JUN-2002
	DEFINITION	AX411211		
	ACCESSION	AX411211		
	VERSION	AX411211.1	GI:21443916	
	KEYWORDS	Homo sapiens (human)		
	SOURCE	Homo sapiens		
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	REFERENCE	I	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
	AUTHORS	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.		
	TITLE	Gene expression profiles in liver cancer		
	JOURNAL	Patent: WO 0229103-A 3858 11-APR-2002;		
	FEATURES	GENE LOGIC INC (US)		
	source	Location/Qualifiers		
		1..2370		
		/organism="Homo sapiens"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:9606"		
		/note="EMBL/GenBank Accession No. Y10032"		
	BASE COUNT	636 a 517 c 513 g 704 t		
	ORIGIN			
	Query Match	100.0%	Score 2370;	DB 6; Length 2370;
	Best Local Similarity	100.0%;	Pred. No. 0;	
	Matches 2370;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
	QY	1	CACGAGGAGCGCTAACGCTCTTCTGCTCCCGCGGTGGTGTGATGACGGTGAACACTGAG	60
	DB	1	CACGAGGAGCGCTAACGCTCTTCTGCTCCCGCGGTGGTGTGATGACGGTGAACACTGAG	60

Qy 61 GCTGCTAAGGGACCCCTCACTTACTCCAGGATGAGGGGCAAGGTGGCAATCTCTACGCT 120
Db |||||
Qy 61 GCTGCTAAGGGACCCCTCACTTACTCCAGGATGAGGGGCAAGGTGGCAATCTCTACGCT 120
Db |||||
Qy 121 TTCATGAAGCAGAGGAGGATGGGTCTGAACGACTTTATTCAGAAGATTGCCAATAACTCC 180
Db |||||
Qy 121 TTCATGAAGCAGAGGAGGATGGGTCTGAACGACTTTATTCAGAAGATTGCCAATAACTCC 180
Db |||||
Qy 181 TATGCATGCAACACCCCTGAAGTTCACTTCTGAGATCTCCCAAGCTCAGAGGCT 240
Db |||||
Qy 181 TATGCATGCAACACCCCTGAAGTTCACTTCTGAGATCTCCCAAGCTCAGAGGCT 240
Db |||||
Qy 241 GAGCTTATGAATGCCAACCCCTCTCTCCACCAAGTCTCTCAGCAATCAACCTTGSC 300
Db |||||
Qy 241 GAGCTTATGAATGCCAACCCCTCTCTCCACCAAGTCTCTCAGCAATCAACCTTGSC 300
Db |||||
Qy 301 CCGTCTGCAATCTCATGCTAAACCATCTGACTTCACTTCTTGAAGCTGATCGAAAG 360
Db |||||
Qy 301 CCGTCTGCAATCTCATGCTAAACCATCTGACTTCACTTCTTGAAGCTGATCGAAAG 360
Db |||||
Qy 361 GGCAGTTTGGAAAGGTTCTCTAGCAAGACCAAGSCAAGAGTGTCTATCGACTC 420
Db |||||
Qy 361 GGCAGTTTGGAAAGGTTCTCTAGCAAGACCAAGSCAAGAGTGTCTATCGACTC 420
Db |||||
Qy 421 AAAGTTTACAGAAAGCAATCCTGAAAGAAAGAGAGAGAGCAATATATGTCGAG 480
Db |||||
Qy 421 AAAGTTTACAGAAAGCAATCCTGAAAGAAAGAGAGAGAGCAATATATGTCGAG 480
Db |||||
Qy 481 CGGAATCTCTGTTGAAGATGTGAAGCAACCTTCTCTGTTGGGCTTCACTTCTCTTC 540
Db |||||
Qy 481 CGGAATCTCTGTTGAAGATGTGAAGCAACCTTCTCTGTTGGGCTTCACTTCTCTTC 540
Db |||||
Qy 541 CAGACTGCTGACAAATGTACTTGTCTTAGACTACATTAATGTTGAGAGTGTCTTAC 600
Db |||||
Qy 541 CAGACTGCTGACAAATGTACTTGTCTTAGACTACATTAATGTTGAGAGTGTCTTAC 600
Db |||||
Qy 601 CATCTCAGAGGAAACCTGCTTCTGGAACCAAGGCTGCTTCTTCTATGCTGCTGAATA 660
Db |||||
Qy 601 CATCTCAGAGGAAACCTGCTTCTGGAACCAAGGCTGCTTCTTCTATGCTGCTGAATA 660
Db |||||
Qy 661 CCCAGTGGCTTGGGCTACCTGCAATCTCTGCAACATCGTTTATAGAGACTTAAACAGAG 720
Db |||||
Qy 661 CCCAGTGGCTTGGGCTACCTGCAATCTCTGCAACATCGTTTATAGAGACTTAAACAGAG 720
Db |||||
Qy 721 AATATTTGCTAGATTACAGGGACACATTTGCTTACTGATTTGCGACTCTGCAAGGAG 780
Db |||||
Qy 721 AATATTTGCTAGATTACAGGGACACATTTGCTTACTGATTTGCGACTCTGCAAGGAG 780
Db |||||
Qy 781 AACATTGAACACACACACACATCCACCTCTCTGCGACGCCGAGTATCTCGACCT 840
Db |||||
Qy 781 AACATTGAACACACACACACATCCACCTCTCTGCGACGCCGAGTATCTCGACCT 840
Db |||||
Qy 841 GAGTGTCTCATAGCAGCCCTTATGACAGGACTGTGACTGTGGTGGCTGGGAGCTGTC 900
Db |||||
Qy 841 GAGTGTCTCATAGCAGCCCTTATGACAGGACTGTGACTGTGGTGGCTGGGAGCTGTC 900
Db |||||
Qy 901 TTGTTATGAGATGCTGATGGCTGCGCCCTTTTATAGCGCAACACAGCTGAATGTAC 960
Db |||||
Qy 901 TTGTTATGAGATGCTGATGGCTGCGCCCTTTTATAGCGCAACACAGCTGAATGTAC 960
Db |||||
Qy 961 GACAACATTTCTGAACAGCCCTCTCCAGCTGAAACCAATATTAACAAATCCGCAAGACAC 1020
Db |||||
Qy 1021 CTCCTGGAGGCCCTCTGACAGAGCAGACCAAGCGCTCGGGCCAGGATGACTTC 1080
Db |||||
Qy 1021 CTCCTGGAGGCCCTCTGACAGAGCAGACCAAGCGCTCGGGCCAGGATGACTTC 1080
Db |||||
Qy 1081 ATGGAGATTAAGATGCTATGCTCTCTCTCTTAACTTAACTGGGATGATCTCAATATAAG 1140
Db |||||
Qy 1081 ATGGAGATTAAGATGCTATGCTCTCTCTCTTAACTTAACTGGGATGATCTCAATATAAG 1140
Db |||||

Qy 1141 AAGATTACTCCCTCTTTAACTCAAAATGTAGTGGGCCCAAGAGCTAGGCACTTTGAC 1200
Db |||||
Qy 1141 AAGATTACTCCCTCTTTAACTCAAAATGTAGTGGGCCCAAGAGCTAGGCACTTTGAC 1200
Db |||||
Qy 1201 CCCAGTTTACGAAAGAGCTGTGCCCAACTCCATTGGCAAGTCCCTGACAGCCTTTC 1260
Db |||||
Qy 1201 CCCAGTTTACGAAAGAGCTGTGCCCAACTCCATTGGCAAGTCCCTGACAGCCTTTC 1260
Db |||||
Qy 1261 GTACAGCCAGCTGCAAGGAAGTGTGCGAGGCTTTCTAGGCTTTTCTATGAGCTTC 1320
Db |||||
Qy 1261 GTACAGCCAGCTGCAAGGAAGTGTGCGAGGCTTTCTAGGCTTTTCTATGAGCTTC 1320
Db |||||
Qy 1321 ACGGACTCTTCTCTCAACCCCTGTAGGCTTTGGTGTAAAGCAATTTATGCTGCTC 1380
Db |||||
Qy 1321 ACGGACTCTTCTCTCAACCCCTGTAGGCTTTGGTGTAAAGCAATTTATGCTGCTC 1380
Db |||||
Qy 1381 CGAATGTTTGTAGTACCTTTTGTGTGAGCGCCGAGCTCAGAGAGCTTTTAAAGAA 1440
Db |||||
Qy 1381 CGAATGTTTGTAGTACCTTTTGTGTGAGCGCCGAGCTCAGAGAGCTTTTAAAGAA 1440
Db |||||
Qy 1441 TTGCGACATCTCTGGAAGCTTAGCAATCTTATTGACACACTGTTCGCTGGAATTTTGA 1500
Db |||||
Qy 1441 TTGCGACATCTCTGGAAGCTTAGCAATCTTATTGACACACTGTTCGCTGGAATTTTGA 1500
Db |||||
Qy 1501 GAGCAGATCTCTCTGAGTGTAGAGCTTTTCAATTTTATTTCTTCCCTCCAGCTG 1560
Db |||||
Qy 1501 GAGCAGATCTCTCTGAGTGTAGAGCTTTTCAATTTTATTTCTTCCCTCCAGCTG 1560
Db |||||
Qy 1561 TGCTATCTCTGAAACAGCGTTAGAGTGGCGCTTAGAGGGAGGAGAGTGTTCATAGA 1620
Db |||||
Qy 1561 TGCTATCTCTGAAACAGCGTTAGAGTGGCGCTTAGAGGGAGGAGAGTGTTCATAGA 1620
Db |||||
Qy 1621 AAGCGGACCTGTTCTTAAAGAGTCTCTGACAGATCTGTGCTGCTGATGACAAATAT 1680
Db |||||
Qy 1621 AAGCGGACCTGTTCTTAAAGAGTCTCTGACAGATCTGTGCTGCTGATGACAAATAT 1680
Db |||||
Qy 1681 TATGAAATGTGCTTTTCTGAAAGAGTGTGTAGTCTCCAAAGCTTTTCTATGCTAG 1740
Db |||||
Qy 1681 TATGAAATGTGCTTTTCTGAAAGAGTGTGTAGTCTCCAAAGCTTTTCTATGCTAG 1740
Db |||||
Qy 1741 TTTGAGTCTTTTATTTTCCCTTCTGAGATATGCTGTGTGACCTGCTGTGCTAGTAT 1800
Db |||||
Qy 1741 TTTGAGTCTTTTATTTTCCCTTCTGAGATATGCTGTGTGACCTGCTGTGCTAGTAT 1800
Db |||||
Qy 1801 GCCTGATCAGAGTGGATTTTGTATGAAGATCAATGTGACACTTCGAGGAGACTACAC 1860
Db |||||
Qy 1801 GCCTGATCAGAGTGGATTTTGTATGAAGATCAATGTGACACTTCGAGGAGACTACAC 1860
Db |||||
Qy 1861 GTGGGACATTTGCTTTCTTCCATATTTGGAAGATAAATTTATGTGTGAGACTTTTGT 1920
Db |||||
Qy 1861 GTGGGACATTTGCTTTTCTTCCATATTTGGAAGATAAATTTATGTGTGAGACTTTTGT 1920
Db |||||
Qy 1921 AAGTACGGTTTAACTAAATTTTATTGAATGGCTTGTGCAATGACTCTGATTCAGATG 1980
Db |||||
Qy 1921 AAGTACGGTTTAACTAAATTTTATTGAATGGCTTGTGCAATGACTCTGATTCAGATG 1980
Db |||||
Qy 1981 CCTAAGAAAGACTGCTGCTACAAATTTCTATTTTATAGAGGCTTTTATGACCA 2040
Db |||||
Qy 1981 CCTAAGAAAGACTGCTGCTACAAATTTCTATTTTATAGAGGCTTTTATGACCA 2040
Db |||||
Qy 2041 ATGCCCGAGTTGTGCTGAGAGCGCTGTGGTGTGTTTCTATGTTTAAATGTCCACTGTA 2100
Db |||||
Qy 2041 ATGCCCGAGTTGTGCTGAGAGCGCTGTGGTGTGTTTCTATGTTTAAATGTCCACTGTA 2100
Db |||||
Qy 2101 AATGGGCAATTTATGCTTTTCTTCTGATGATGATGATGATGATGATGATGATGAT 2160
Db |||||
Qy 2101 AATGGGCAATTTATGCTTTTCTTCTGATGATGATGATGATGATGATGATGATGAT 2160
Db |||||
Qy 2161 AACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db |||||
Qy 2161 AACCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db |||||
Qy 2221 TAAACCAACCTTTTAAATGCTACTGTAATTAACATGGTGTATTAACATGCTTCCCTC 2280
Db |||||

```

Db      2221  TAAACACCATTTTAACTACTGTAAATTAACATGGTTATATACGATCAATCCCTCCCTC 2280
QY      2281  ATCCCATCACACAACTTTTGTGTGATGAATAAAGTATTTGGTTTGCATAAAGACCTT 2340
Db      2281  ATCCCATCACACAACTTTTGTGTGATGAATAAAGTATTTGGTTTGCATAAAGACCTT 2340
QY      2341  GAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2370
Db      2341  GAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2370

RESULT 5
LOCUS   HSRNATPK                2370 bp    mRNA    linear    PRI 16-MAY-1997
DEFINITION H.sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION Y10032
VERSION   Y10032.1 GI:1834510
KEYWORDS  serine/threonine protein kinase; sgk gene.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS   Waldegger, S., Barth, P., Raber, G. and Lang, F.
TITLE     Cloning and characterization of a putative human serine/threonine
           protein kinase transcriptionally modified during anisotonic and
           isotonic alterations of cell volume
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 94 (9), 4440-4445 (1997)
MEDLINE   97272242
PUBMED    9114008
REFERENCE 2 (bases 1 to 2370)
AUTHORS   Waldegger, S.
TITLE     Direct Submission
JOURNAL   Submitted (11-DEC-1996) S. Waldegger, University of Tuebingen,
           Physiology I, Gmelinstr. 5, D-72076 Tuebingen, FRG
FEATURES
     source
         1..2370
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="adult"
             /gene="sgk"
             /gene="sgk"
             /codon_start=1
             /product="serine/threonine protein kinase"
             /protein_id="CAA71138.1"
             /db_xref="GI:1834511"
             /translation="MTVTEAAKGLTYSRMGMVAILIAFMKORRMGLNDFIOKIAN
             NSYACKHPVQILKISQPEPELMNPPSPSPSQINLGPSSNPAKSDHFLLK
             VIGFSGFKVLARKAEVFAVKVQKALKKKKEKHINSRNLKKNVHPFLV
             GHFSEQADKLYFVLOYINGGEVFLHQRERCFLEPRAREYAAEIASALGYLSLNI
             VYDLPKLENLDSQGHVLYDFGLCKENIEHNSSTFCGTPPEYLAPEVLHQPDR
             TVDWCLGAVLYEMLYGLPFFYSRNTAEYDNLNKLQKPNITNSRHLLGGLQK
             DRTRKLGADKDMELKSHVFFSLINWDLINKKITPPFNPVSGPNELRHDFPEFEE
             PVPNSIGKSPDSVLYTASVKERAELFCSYAPDPSFL"
BASE COUNT 636 a 517 c 513 g 704 t
ORIGIN
Query Match 100.0%; Score 2370; DB 9; Length 2370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CACAGAGGCGGTACGCTTTCTCTCTCCCGCGGTGGTGATGACGGTGAACACTGAG 60
Db      1  CACAGAGGCGGTAAACGCTTTCTCTCTCCCGCGGTGGTGATGACGGTGAACACTGAG 60
QY      61  GCTGCTAAGGCGACCTCACTTACTCCAGGATCAGGGGATGCTGGCAATTCCTCATCGCT 120
Db      61  GCTGCTAAGGCGACCTCACTTACTCCAGGATCAGGGGATGCTGGCAATTCCTCATCGCT 120

```

```

QY      121  TTCAATGAAGCAGAGGAGGATGGCTCTGAACGACTTTTATTCAGAGAGATTGCTCAATACCTC 180
Db      121  TTCAATGAAGCAGAGGAGGATGGCTCTGAACGACTTTTATTCAGAGAGATTGCTCAATACCTC 180
QY      181  TATGTCATGCAACACACCTGAAAGTTTCAGTCCATCTTGAAGATCTCCCAAGCTCAGAGAGCT 240
Db      181  TATGTCATGCAACACACCTGAAAGTTTCAGTCCATCTTGAAGATCTCCCAAGCTCAGAGAGCT 240
QY      241  GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCCCTTCCTCAGCAAAATCAACCTTGCC 300
Db      241  GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCCCTTCCTCAGCAAAATCAACCTTGCC 300
QY      301  CCGTCGTCCTCAATCCTCATGCTAAACCATCTGACTTTTCACTTCTTGAAGTGTATGGAAG 360
Db      301  CCGTCGTCCTCAATCCTCATGCTAAACCATCTGACTTTTCACTTCTTGAAGTGTATGGAAG 360
QY      361  GGCAGTTTTCGAAAGTTTCTTAGCAAGACACACAGGAGAGAAAGTGTTCATGAGTTC 420
Db      361  GGCAGTTTTCGAAAGTTTCTTAGCAAGACACACAGGAGAGAAAGTGTTCATGAGTTC 420
QY      421  AAAGTTTACAGAGAGAGCAATCCTCAAAAGAAAGAGAGAGAGATATATATGTCGAG 480
Db      421  AAAGTTTACAGAGAGAGCAATCCTCAAAAGAAAGAGAGAGAGATATATATGTCGAG 480
QY      481  CGGAATGTTCTGTGTAAGAATGTGAAGCACCCCTTTCTGCTGGGCGCTTCACTTCTTCTTC 540
Db      481  CGGAATGTTCTGTGTAAGAATGTGAAGCACCCCTTTCTGCTGGGCGCTTCACTTCTTCTTC 540
QY      541  CAGACTGCTGACAAATGTACTTGTCTAGACTACATTAATGCTGAGAGATGTCTTCTAC 600
Db      541  CAGACTGCTGACAAATGTACTTGTCTAGACTACATTAATGCTGAGAGATGTCTTCTAC 600
QY      601  CATCTCCAGAGGAAACCTGCTCTCGGAACACAGGCGCTGTTCTTATGTATGTCGTAATA 660
Db      601  CATCTCCAGAGGAAACCTGCTCTCGGAACACAGGCGCTGTTCTTATGTATGTCGTAATA 660
QY      661  GCCAGTGCCTTGGGCTACCTGCTCTAGACTGATTAATGCTGAGAGATGTCTTCTAC 720
Db      661  GCCAGTGCCTTGGGCTACCTGCTCTAGACTGATTAATGCTGAGAGATGTCTTCTAC 720
QY      721  AATATTTGCTAGATTTCACAGGACACATTCCTTACTGATTTTCGAGACTCTGCAAGAG 780
Db      721  AATATTTGCTAGATTTCACAGGACACATTCCTTACTGATTTTCGAGACTCTGCAAGAG 780
QY      781  AACATTTGAACACACACACACATCCACCTTCTGTGGCAGCGCGGAGTATCTGCAACT 840
Db      781  AACATTTGAACACACACACACATCCACCTTCTGTGGCAGCGCGGAGTATCTGCAACT 840
QY      841  GAGTGTCTTCATAAGCAGCCTTATGACAGGACTGTGACTGGTGGTGGTGGGAGTGTGTC 900
Db      841  GAGTGTCTTCATAAGCAGCCTTATGACAGGACTGTGACTGGTGGTGGTGGGAGTGTGTC 900
QY      901  TTGATGAGATGCTGTATGGCTCCGCCCTTTTATAGCGCAACACAGCTGAAATGTATC 960
Db      901  TTGATGAGATGCTGTATGGCTCCGCCCTTTTATAGCGCAACACAGCTGAAATGTATC 960
QY      961  GACAACTTCGAAACAGCCTCTCCAGCTGAACCAATATTAACAATTCGCAAGACAC 1020
Db      961  GACAACTTCGAAACAGCCTCTCCAGCTGAACCAATATTAACAATTCGCAAGACAC 1020
QY      1021  CTCCTGGAGGCGCTTCCTGCAGAGAGACAGGACAAAGCGGCTCGGGGCCAAGAGTATCTTC 1080
Db      1021  CTCCTGGAGGCGCTTCCTGCAGAGAGACAGGACAAAGCGGCTCGGGGCCAAGAGTATCTTC 1080
QY      1081  ATGAGAGATTAGAGATCATGCTTCTTCTCTTAACTTAACTGGGATGATCTCATTAAG 1140
Db      1081  ATGAGAGATTAGAGATCATGCTTCTTCTCTTAACTTAACTGGGATGATCTCATTAAG 1140
QY      1141  AAGATTACTCCCTCTTTTAAACCAAAATGTAGTGGGCGCCACAGCTACCGCACTTTGAC 1200
Db      1141  AAGATTACTCCCTCTTTTAAACCAAAATGTAGTGGGCGCCACAGCTACCGCACTTTGAC 1200

```

Qy	1201	CGCAGCTTACCGAAGAGCCTGTCGCCCAACTCCATTGGCAAGTCCCTCGACAGAGTCTCCTC	1260
Db	1201	CCCGAGTTTTACCGAAGAGAGCCTGTCGCCCAACTCCATTGGCAAGTCCCTCGACAGAGTCTCCTC	1260
Qy	1261	GTCCACAGCCAGCGTCAAGGAAGCTGCCGAGGCTTTCCCTAGGCTTTTCCCTATGCCGCTGCC	1320
Db	1261	GTCCACAGCCAGCGTCANAGGAAGCTGCCGAGGCTTTCCCTAGGCTTTTCCCTATGCCGCTGCC	1320
Qy	1321	ACGGACTCTTTCCCTCTGAACCCCTGTGTAGGCTTGGTTTTAAAGGATTTTATGTGTGTTC	1380
Db	1321	ACGGACTCTTTCCCTCTGAACCCCTGTGTAGGCTTGGTTTTAAAGGATTTTATGTGTGTTC	1380
Qy	1381	CGAATGTTTTAGTTAGCTTTGGTGGAGCCGCCAGCTGACAGGACATCTTTACAAGAGAA	1440
Db	1381	CGAATGTTTTAGTTAGCTTTGGTGGAGCCGCCAGCTGACAGGACATCTTTACAAGAGAA	1440
Qy	1441	TTTGCACATCTCTGGAGAGCTTAGCAATCTTATGACACTGTTCGCTGGAATTTTTTGAA	1500
Db	1441	TTTGCACATCTCTGSAAGCTTAGCAATCTTATGACACTGTTCGCTGGAATTTTTTGAA	1500
Qy	1501	GAGCACATCTCCTCAGTGAGCTCATGAGGTTTTCAATTTTATCTTCCTCCAAACGTGG	1560
Db	1501	GAGCACATCTCCTCAGTGAGCTCATGAGGTTTTCAATTTTATCTTCCTCCAAACGTGG	1560
Qy	1561	TGCTATCTCTGAACAGAGCTTAGAGTCCGCGCTTAGACGAGGAGGAGTTCCTTAGA	1620
Db	1561	TGCTATCTCTGAACAGAGGCTTAGAGTCCGCGCTTAGACGAGGAGGAGTTCCTTAGA	1620
Qy	1621	AAGCGGACCTGTTCTTAAAAAGCTCCTCGACATCTGCTGGGCTGTGATGACGAATAT	1680
Db	1621	AAGCGGACCTGTTCTTAAAAAGCTCCTCGACATCTGCTGGGCTGTGATGACGAATAT	1680
Qy	1681	TATGAATGTGCCFTTTCTCAAGAGATGTGTTAGCTCCAAAGCTTTTCTTATCCAGTGTG	1740
Db	1681	TATGAATGTGCCFTTTCTCAAGAGATGTGTTAGCTCCAAAGCTTTTCTTATCCAGTGTG	1740
Qy	1741	TTTCAGTCTTTATTTTCCCTGTGGATATGCTGTGGAACCGCTGTGAGGTGGTAT	1800
Db	1741	TTTCAGTCTTTATTTTCCCTGTGGATATGCTGTGGAACCGCTGTGAGGTGGTAT	1800
Qy	1801	GCCTGATCAGAGATGGATTTGTTTATAAGCATCAATGTGACACTTCGAGGACACTACAAC	1860
Db	1801	GCCTGATCAGAGATGGATTTGTTTATAAGCATCAATGTGACACTTCGAGGACACTACAAC	1860
Qy	1861	GTGGGACATTTGTTTCTCCATATTTGGAGATATAATTTATGTGTAGCTTTTGTGT	1920
Db	1861	GTGGGACATTTGTTTCTCCATATTTGGAGATATAATTTATGTGTAGCTTTTGTGT	1920
Qy	1921	AAGATACGGTTAATAACTAAAAATTTATGAAATGGCTTGCATATGACTCGTATTCAGATG	1980
Db	1921	AAGATACGGTTAATAACTAAAAATTTATGAAATGGCTTGCATATGACTCGTATTCAGATG	1980
Qy	1981	CCTTAACAAAGCATTGCTGCTACAAATATTTCTATTTTAGAAGGGTTTTATGAGACCA	2040
Db	1981	CCTTAAGAAAGCATTGCTGCTACAAATATTTCTATTTTAGAAGGGTTTTATGAGACCA	2040
Qy	2041	ATGCCCCAGTTGTCCAGTCAGAGCGGTTGGTGTTCATTTGTTTAAATATGTCACCTGTAA	2100
Db	2041	ATGCCCCAGTTGTCCAGTCAGAGCGGTTGGTGTTCATTTGTTTAAATATGTCACCTGTAA	2100
Qy	2101	AATGGGATATTTATGTTTTTTTTTTTTTGTGATCTCCGATAATTTGATGATTCGATAAAG	2160
Db	2101	AATGGGATATTTATGTTTTTTTTTTTTTGTGATCTCCGATAATTTGATGATTCGATAAAG	2160
Qy	2161	AACGTCTACATTTGGGTTATACACTAGTATATTTAAACTTACAGGCTTATTTGTAATG	2220
Db	2161	AACGTCTACATTTGGGTTATACACTAGTATATTTAAACTTACAGGCTTATTTGTAATG	2220
Qy	2221	TAAACCAACATTTTAAAGTACTGTAATTAACATGGTTATATAACGTACATCTCTCCCTC	2280
Db	2221	TAAACCAACATTTTAAAGTACTGTAATTAACATGGTTATATAACGTACATCTCTCCCTC	2280
Qy	2281	ATCCCCATCACACAACCTTTTTTTTGTGTGTGATAAACTGATTTTGGTTTTGCAATAAAACCTT	2340

```

|||||
2281 ATCCCATACACAACTTTTGTGTCGTATTAACCTGATTGGTTTGCAATAAACCCT 2340
|||||
Qy 2341 GAAAAATAAAAAAAAAAAAAAAAAAAAAA 2370
|||||
Db 2341 GAAAAATAAAAAAAAAAAAAAAAAAAAAA 2370
|||||

RESULT 6
AF153609      2382 bp   mRNA       linear    PKI 28-JUN-1999
LOCUS        Homo sapiens serine/threonine protein kinase sgk mRNA, complete
cgs.
ACCESSION    AF153609
VERSION      AF153609.1 GI:5231142
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 2382)
AUTHORS     Kim,M.K., Kim,Y.H., Suh,J.M., Lee,H.M., Chung,H.J., Sohn,M.Y.,
Hwang,S.Y., Im,S.U., Jung,E.J. and Kim,J.C.
TITLE       A catalogue of genes in the human dermal papilla cells as
identified by expressed sequence tags
JOURNAL
REFERENCE    Unpublished
AUTHORS
2 (bases 1 to 2382)
Kim,M.K., Kim,Y.H., Suh,J.M., Lee,H.M., Chung,H.J., Sohn,M.Y.,
Hwang,S.Y., Im,S.U., Jung,E.J. and Kim,J.C.
DIRECT SUBMISSION
TITLE       Submitted (24-MAY-1999) Immunology, Kyungpook National University,
SCHOOL OF Medicine, 101 Donggong, Jung Gu, Taegu, Taeju 700-422,
South Korea
FEATURES
source       1. 2382
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /cell_line="primary culture"
              /cell_type="hair dermal papilla"
              43..1338
              /codon_start=1
              /product="serine/threonine protein kinase sk"
              /protein_id="AAD41091.1"
              /db_xref="GI:5231143"
              /translation="MTVKTBAKQTITYSRMGMVAIIAPMKQRKGILNDFLTKIAN
NSYACKFGEVQSILIKTSOPEPELMANPSPPPSPSQQINIGPSNNHPAKPSDFHLPL
VIGSGTFGVLLARHKAEEFYAVKVLOKKAILKKKEKHIMSENVNLKNVKHFLPV
GLHFSPOTAKLVFDLYINGGELYFHLORECFLEPRARFYAAETASALGYLHSLNI
VYRLDKPENILLSSGHIVLTDFGLCKENIEHNSTTSTFCGTPEYLAPAEVLHKOPYDR
TDWMWCLAGHYEMLYGLPFYPSRNTAEVDYNLNKPQLKPNITNSARHLLGLLQK
DTRKLGAQUDFMELKSHVPFSLLNWDDLNKKITPPFPNPVSVPNDLRHFDPTFEET
PVPNSIGKSPDSVLVTASVRAEAFLGFSAVPTDSDL"

BASE COUNT   640 a   519 c   513 g   710 t
ORIGIN

Query Match          98.2%; Score 2328.2; DB 9; Length 2382;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2355; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

Qy 8 GAGCGCTAACCTCTTCTGTCTCCCCCGGTGGTGATGACGGTCAAACACTGAGCCTGCTA 67
|||
Db 8 GAGCGCTAACCTCTTCTGTCTCCCCCGGTGGTGATGACGGTCAAACACTGAGCCTGCTA 67
|||

Qy 68 AGGCACCCCTCACTTACTCYAGGATGAGGGGCATGGTGCCAATTTCTCATGGCTTTCATGA 127
|||
Db 68 AGGCACCCCTCACTTACTCCAGGATGAGGGGCATGGTGCCAATTTCTCATGGCTTTCATGA 127
|||

Qy 128 ACAGAGAGAGATGGGCTGAACGACTTTATTTCAGAAGATGCCAATAACTCCTATGCAT 187
|||
Db 128 ACAGAGAGAGATGGGCTGAACGACTTTATTTCAGAAGATGGCCAATAACTCCTATGCAT 187
|||

Qy 188 GCACAACACCCCTGAAGCTTCAGTCTCCATCTTGAAGATCTCCCAACCTCAGCAGCCTGAGCTTA 247

```

188 GCACACACCTGAAGTTCAGTCCATCTGAAGATCTCCCAACCTCAGGAGCCTGAGCTTA 247
248 TGAATGCCAACCTTCTCCTCCACCAAGTCTCTTCAGCAATCAACCTTGGCCCGTGT 307
249 TGAATGCCAACCTTCTCCTCCACCAAGTCTCTTCAGCAATCAACCTTGGCCCGTGT 307
308 CCAATCCTCATCTAAACCATCTGACTTTCACCTTCTGAAAGTGATCGAAAGGCGAGTT 367
309 CCAATCCTCATCTAAACCATCTGACTTTCACCTTCTGAAAGTGATCGAAAGGCGAGTT 367
368 TTGGAAAGGTTCTTCTAGCAACACACAGGCGAGAGAGTGTCTATGCAAGTCAAGTTT 427
369 TTGGAAAGGTTCTTCTAGCAACACACAGGCGAGAGAGTGTCTATGCAAGTCAAGTTT 427
428 TACAGAAGAAACCAATCTCTGAAAGAGAGAGAGAGATATATGTCGAGGCGGAATG 487
429 TACAGAAGAAACCAATCTCTGAAAGAGAGAGAGAGATATATGTCGAGGCGGAATG 487
488 TTTCTGTTGAAGAAATGTGAAGCACCTTTCTCTGGTGGGCTTTCACCTTCTTTCAGACTG 547
489 TTTCTGTTGAAGAAATGTGAAGCACCTTTCTCTGGTGGGCTTTCACCTTCTTTCAGACTG 547
548 CTGACAAATGTGACTTGTCTCTAGACTACATTAATGTTGGAGTGTCTTACCAATCTCC 607
549 CTGACAAATGTGACTTGTCTCTAGACTACATTAATGTTGGAGTGTCTTACCAATCTCC 607
608 AGAGGAACGCTGCTTCTGGAACACACCGGCTCTTCTATGCTGCTGAAATAGCAGTG 667
609 AGAGGAACGCTGCTTCTGGAACACACCGGCTCTTCTATGCTGCTGAAATAGCAGTG 667
668 CTTTGGGCTACCTGATCTGACAGTGTGTTTATAGACTTTAAACACAGAGATATTT 727
669 CTTTGGGCTACCTGATCTGACAGTGTGTTTATAGACTTTAAACACAGAGATATTT 727
728 TCTAGATTCACAGGACACATGTCCTTACTGATTTCTGGACTCTGCAAGGAGACATTTG 787
729 TCTAGATTCACAGGACACATGTCCTTACTGATTTCTGGACTCTGCAAGGAGACATTTG 787
788 AACACACAGACACACATCCACCTTCTGTGACCCGCGAGTATCTGCACCTGAGGTGC 847
789 AACACACAGACACACATCCACCTTCTGTGACCCGCGAGTATCTGCACCTGAGGTGC 847
848 TTCATAGCAGCCTTATGACAGGACTGTGACTGTGTGCTGGGAGCTGTCTGTATG 907
849 TTCATAGCAGCCTTATGACAGGACTGTGACTGTGTGCTGGGAGCTGTCTGTATG 907
908 AGATGCTGTATGGCTGCGGCTTTTATAGCCGAAACACAGCTGAATGTACGACAACA 967
909 AGATGCTGTATGGCTGCGGCTTTTATAGCCGAAACACAGCTGAATGTACGACAACA 967
968 TTCGACACAGCCTTCTCAGCTGAACCAATATTTACAAATTCGCGAAGACACCTTCTGG 1027
969 TTCGACACAGCCTTCTCAGCTGAACCAATATTTACAAATTCGCGAAGACACCTTCTGG 1027
1028 AGGGCTCTCTCAGAGGACAGGACAAAGCGCTGCGGGCCCAAGGATGACTTCATGGAGA 1087
1029 AGGGCTCTCTCAGAGGACAGGACAAAGCGCTGCGGGCCCAAGGATGACTTCATGGAGA 1087
1088 TTAAGAGTCATGCTTCT 1147
1089 TTAAGAGTCATGCTTCT 1147
1148 CTCGCCCTTTTAAACCAATGTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1207
1149 CTCGCCCTTTTAAACCAATGTAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1207
1208 TTACCGAAGAGCCTGTCTCCCAACTCCATTTGCAAGTCCCTGACAGGCTCTCTGTCACAG 1267
1209 TTACCGAAGAGCCTGTCTCCCAACTCCATTTGCAAGTCCCTGACAGGCTCTCTGTCACAG 1267
1268 CCAGCGTCAAGAGAGCTGCGGAGGCTTTCTAGGCTTTTCTATGCGGCTCCCAAGGACT 1327

1269 CCAGCGTCAAGAGAGCTGCGGAGGCTTTCTTAGGCTTTTCTATGCGGCTCCCAAGGACT 1327
1328 CTTTCTCTGAACCCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCTCAATGT 1387
1329 CTTTCTCTGAACCCCTGTAGGCTTGGTTTAAAGGATTTTATGTGTGTTTCTCAATGT 1387
1388 TTTAGTTAGCCTTTTGGTGGAGCCGAGCTGACAGGACATCTTACAGAGAAATTTGAC 1447
1389 TTTAGTTAGCCTTTTGGTGGAGCCGAGCTGACAGGACATCTTACAGAGAAATTTGAC 1447
1448 ATCTCTGGAAGCTTAGCAATCTTATGACACTGTTCGCTGGAAGCTTTTGAAGAGCAC 1506
1449 ATCTCTGGAAGCTTAGCAATCTTATGACACTGTTCGCTGGAAGCTTTTGAAGAGCAC 1507
1507 ATCTCTCTCAGTGAGCTCATGAGGTTTTCATTTTATTTCTCTTCCCAAGCTGTTAT 1566
1508 ATCTCTCTCAGTGAGCTCATGAGGTTTTCATTTTATTTCTCTTCCCAAGCTGTTAT 1567
1567 CTCTGAAACGAGCGTTAGAGTCCGCCCTTAGACGAGGAGGAGTTTCTTTAGAAAGGG 1626
1568 CTCTGAAACGAGCGTTAGAGTCCGCCCTTAGACGAGGAGGAGTTTCTTTAGAAAGGG 1627
1627 AC-CTGTTCTTAAAGGCTCTCTGACATCTCTGAGTGTGCGCTGTGATGACGAAATATATGA 1685
1628 ACCTGTTCTTAAAGGCTCTCTGACATCTCTGAGTGTGCGCTGTGATGACGAAATATATGA 1687
1686 AATGTGCTTTCTGAGAGAGTGTGTTAGCTCCAAAGCTTTCTATCTCCAGTGTATCA 1745
1687 AATGTGCTTTCTGAGAGAGTGTGTTAGCTCCAAAGCTTTCTATCTCCAGTGTATCA 1747
1746 GTTCTTTATTTTCCCTTGTGATATGCTGTGAACCTGCTGTGAGTGTGTATGCTG 1805
1747 GTTCTTTATTTTCCCTTGTGATATGCTGTGAACCTGCTGTGAGTGTGTATGCTG 1807
1806 ATCAGAGATGATTTTGTATTAAGCATCAATGTGACACTTCGAGGACACTACAAGTGG 1865
1807 ATCAGAGATGATTTTGTATTAAGCATCAATGTGACACTTCGAGGACACTACAAGTGG 1867
1866 ACATTTGTTGTTTCTTCCATATTTTGAAGATAAATTTATGCTAGACTTTTCTTGAAGAT 1925
1867 ACATTTGTTGTTTCTTCCATATTTTGAAGATAAATTTATGCTAGACTTTTCTTGAAGAT 1927
1926 ACGGTTAATACTAAATTTTATGAAATGCTTTCGCAATGACTCGTATTCAGATGCTTAA 1985
1927 ACGGTTAATACTAAATTTTATGAAATGCTTTCGCAATGACTCGTATTCAGATGCTTAA 1987
1986 AGAAGCATTCCTGCTACAAATATTTCTATTTTGAAGAGGCTTTTATGAGCAATGCTC 2045
1987 AGAAGCATTCCTGCTACAAATATTTCTATTTTGAAGAGGCTTTTATGAGCAATGCTC 2047
2046 CCAGTTGCTCAGTCAAGCGCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2105
2047 CCAGTTGCTCAGTCAAGCGCTTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2107
2106 GCATTTATTTGTTTTTTTTTTCATTTCTGATAATTTGATATGATGATGATGATGATGAT 2165
2107 GCATTTATTTGTTTTTTTTTTCATTTCTGATAATTTGATATGATGATGATGATGATGAT 2167
2166 CTGTACATTTGGTTTATACACTAGTATATTTAAACTTACAGGCTTTTGTATGTAATGTAAC 2225
2167 CTGTACATTTGGTTTATACACTAGTATATTTAAACTTACAGGCTTTTGTATGTAATGTAAC 2227
2226 CACCATTTAATGTACTGTAATTAACATGTTTATATAGTACAACTCTTCCCTCATCCC 2285
2287 CACCATTTAATGTACTGTAATTAACATGTTTATATAGTACAACTCTTCCCTCATCCC 2287
2288 ATCACACACTTTTTTGTGTGTGATAAAGTATTTTGGTTTGTGATAAAGTATTTGATAA 2345
2289 ATCACACACTTTTTTGTGTGTGATAAAGTATTTTGGTTTGTGATAAAGTATTTGATAA 2347
2346 AAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2370
2347 AAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2372


```

RESULT 7
BC001263
LOCUS BC001263
DEFINITION Homo sapiens, serum/glucocorticoid regulated kinase, clone MGC:5030
IMAGE:3459056, mRNA, complete cds.
ACCESSION BC001263
VERSION BC001263.1
KEYWORDS GI:12654838
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2382)
Direct Submission
Submitted (11-DEC-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcapps-femail.nih.gov
Tissue Procurement: AFCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 4 Row: a Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5032090.
FEATURES
source
1. .2382
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:6446"
/db_xref="taxon:9606"
/clone="MGC:5030 IMAGE:3459056"
/tissue_type="Cervix, carcinoma"
/clone_lib="NIH_MGC_12"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
43. .1338
/codon_start=1
/product="serum/glucocorticoid regulated kinase"
/protein_id="AA001263.1"
/db_xref="GI:12654838"
/translation="MTVKEAAGTLYSRMRQWVAIIAFMKRRMGLNDFIQIAN
NSYAKHPVQSILKISQEPPELMNANPPSPSQINLGPSSNHPAKPSDFHFLK
VIGKSGFQVLLARHKAEEFYAVKVLQKKAILKKEKHIMSERVLLKVKHPLV
GLHFSQADKLVFLVDYINGGELFYHLORECFLEPRARFYAAETASALGYLHSLNI
VYRDLEPENLLDSQGHVLTDFGLCKENIEHNSSTSTCGTPEVLAPELVHKOYDR
TVDWCLGAVLYEMLKGLVFPFYSRNTAEYDNLNPKPLQKPNITNSARHLEGLLQK
DTRKLGAKDMEIKSHVFSLNDDLLINKITPPFNPNVSGPNDLRHDFPEFTEE
PVFNSIGKSPDVLVTAASKVEAAEALFSGYAPPTDSFL"
BASE COUNT 642 a 519 c 513 g 708 t
ORIGIN
Query Match 97.78; Score 2316.2; DB 9; Length 2382;
Best Local Similarity 99.56; Pred. No. 0;
Matches 2354; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 8 GAGCGCTAACGCTCTTTCTGCTCCCGCGGTGGTGATGACGGTGAAGAACTGAGCGCTGCTA 67
|||||
DB 8 GAGCGCTAACGCTCTTTCTGCTCCCGCGGTGGTGATGACGGTGAAGAACTGAGCGCTGCTA 67
|||||

```


1148 CTCCCTTTTACCCAAATGTGAGTGGCCCAACGAGCTACGGCACTTTGACCCCGAGT 1207
 1148 CTCCCTTTTACCCAAATGTGAGTGGCCCAACGAGCTACGGCACTTTGACCCCGAGT 1207
 1208 TTACGAGAGGCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCCCTGTCACAG 1267
 1208 TTACGAGAGGCTGTCCCAACTCCATTGGCAAGTCCCTGACAGCGTCCCTGTCACAG 1267
 1268 CCAGCGTCAAGGAGCTGCGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGAGCT 1327
 1268 CCAGCGTCAAGGAGCTGCGAGGCTTCCCTAGGCTTTTCCCTATGCGCTCCACGAGCT 1327
 1328 CTTTCTCTGAACCTCTTAGGCGCTTGGTTTAAAGGATTTATGTGTGTTCCGAAATGT 1387
 1328 CTTTCTCTGAACCTCTTAGGCGCTTGGTTTAAAGGATTTATGTGTGTTCCGAAATGT 1387
 1388 TTTAGTTAGCTTTTGGTGAGCGCGCAGCTGACAGGACATCTTACAAGAGATTTGACAC 1447
 1388 TTTAGTTAGCTTTTGGTGAGCGCGCAGCTGACAGGACATCTTACAAGAGATTTGACAC 1447
 1448 ATCTCTGAGAGCTTAGCAATCTTATGCAACACTGTTCCGCGGAA-TTTTGTGAAGAGCAC 1506
 1448 ATCTCTGAGAGCTTAGCAATCTTATGCAACACTGTTCCGCTGGAAGCTTTTGAAGAGCAC 1507
 1507 ATCTCTCTGAGTGTGAGCTTGTGAGCTTTTATTTTATCTTCTTCCCAAGCTGGTGTAT 1566
 1508 ATCTCTCTGAGTGTGAGCTTGTGAGCTTTTATTTTATCTTCTTCCCAAGCTGGTGTAT 1567
 1567 CTCTGAAACGAGCTTAGAGTGCGGCTTAGAGCGGAGGAGGAGGATTTGTTAGAAAGCGG 1626
 1568 CTCTGAAACGAGCTTAGAGTGCGGCTTAGAGCGGAGGAGGAGGATTTGTTAGAAAGCGG 1627
 1627 AC-CTGTCTTAAAGAGCTCTCTGCGAGATCTCTGCGGCTGTGATGACGAAATATTATGA 1685
 1628 ACCTGTCTTAAAGAGCTCTCTGCGAGATCTCTGCGGCTGTGATGACGAAATATTATGA 1687
 1686 AATGTGCTTTCTGAGAGATTTGTTAGCTTCAAGAGCTTTTCTATCGCAGTGTTCAT 1745
 1688 AATGTGCTTTCTGAGAGATTTGTTAGCTTCAAGAGCTTTTCTATCGCAGTGTTCAT 1747
 1746 GTTCTTATTTTCCCTTGTGATGATGCTGTGAAAGCTGCTGAGTGAGTGATGCTCCG 1805
 1748 GTTCTTATTTTCCCTTGTGATGATGCTGTGAAAGCTGCTGAGTGAGTGATGCTCCG 1807
 1806 ATCAGAGATGATTTGTTTAAAGCAATGATGACACTTTCGAGGACACTTACACCTGGG 1865
 1808 ATCAGAGATGATTTGTTTAAAGCAATGATGACACTTTCGAGGACACTTACACCTGGG 1867
 1866 ACATTGTTGTTCTTCCATATTTGGAAGATAATTTATGCTGTAGACTTTTGTGTAAGAT 1925
 1868 ACATTGTTGTTCTTCCATATTTGGAAGATAATTTATGCTGTAGACTTTTGTGTAAGAT 1927
 1926 ACGGTTTAACTAAATTTATTAAGATGCTTTGCAATGACCTGATTCAGATGCTTAA 1985
 1928 ACGGTTTAACTAAATTTATTAAGATGCTTTGCAATGACCTGATTCAGATGCTTAA 1987
 1986 AGAAGCAATGCTGCTCAATATTTCTTATTTTAGAAGGTTTATGAGCACTGATGCC 2045
 1988 AGAAGCAATGCTGCTCAATATTTCTTATTTTAGAAGGTTTATGAGCACTGATGCC 2047
 2046 CCAGTGTGAGTCAAGCGCTGCTGTTTTCATGTTTAAATGTCACCTGTAAATGG 2105
 2048 CCAGTGTGAGTCAAGCGCTGCTGTTTTCATGTTTAAATGTCACCTGTAAATGG 2107
 2106 GCATATTTATGTTTCTTCTGATTCCTGATTAATTTATGATGATTTATTAAGACGT 2165
 2108 GCATATTTATG-TTTTCTTCTGATTCCTGATTAATTTATGATGATTTATTAAGACGT 2166
 2166 CTGTACATGCTTATACACTAGTATTTTAACTTACAGGCTTATTTGTAATGTAAC 2225
 2167 CTGTACATGCTTATACACTAGTATTTTAACTTACAGGCTTATTTGTAATGTAAC 2226
 2226 CACCATTATGACTGTAATTAACATGGTGTATTAATACGTACAACTTCCCTCCCTCATCC 2285

2227 CACCATTATGACTGTATTAACATGGTTAATACATGATTAATACATGATTAATACATCC 2286
 2286 ATACACAACTTTTGTGTGTGATAACTGATTTTGGTTTGCATTAATAAACCTTGAATA 2345
 2287 ATACACAACTTTTGTGTGTGATAACTGATTTTGGTTTGCATTAATAAACCTTGAATA 2346
 2346 ATAAAAAATAAAAAAATAAAAAA 2370
 2347 ATATTAATAAAAAAATAAAAAA 2371
 RESULT 8
 AX553549
 LOCUS AX553549
 DEFINITION Sequence 1 from Patent WO02074987.
 ACCESSION AX553549
 VERSION AX553549.1 GI:25897549
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1
 Lang, F., Busjahn, A. and Luft, F.C.
 Quantitative diagnostic analysis of hypertonia
 Patent: WO 02074987-A 1 26-SEP-2002;
 Lang, Florian (DE)
 Location/Qualifiers
 1..2354
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="Laxson:9606"
 43..1338
 /note="unnamed protein product"
 /cdoon_start=1
 /protein_id="CAD58123.1"
 /db_xref="GI:25897550"
 /translation="MTVTEAAKGTLTYSRMGRMVAILAFMKORRMGLNDFIOLKIAN
 NSYACKHPEVOSILKIISOPELVANANPPSPSOOLNPPSSNHPHAKPSDFHFLK
 VIKGSGKVLARHAKAEVEYAVKLOKKAILKKKEKHIMSENVLLAKNVKHPFLV
 GLHFSQPADKLYLVLDYINGELFYHLORERCPLERARFYAEIASALCYLHSLNI
 YRDMLKPNILDDSGHIVLDFOLCKENIEHNTSTTTCGTPEYLAPEVHLKQPTDR
 VTDMCLGAVILEMYLGLPFLYSRNTEMDYNILNKLPLKPNITNSARHLDELLOK
 DTRKRLGAKDDFMEIKSHVFSFLINMDOLINKTKITPPENPNVSGPNOLRHFPEFTEE
 PVPNSIGKSPDSVLVTASVKEAAEAFLOFSYAPPTDSFL"
 762
 variation
 /note="1. SNP (C in T), stumme Mutation, d. h. belde
 Versionen des SNPs resultieren in der AMINO ACID Asp in
 der AMINO ACID-Position 240"
 BASE COUNT 613 a 518 c 513 g 710 t
 ORIGIN
 Query Match 97.6%; Score 2312.6; DB 6; Length 2354;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2337; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
 QY 8 GAGCGCTAACGCTCTTCTGTCTCCCGCGGTGTGTCATGACGGTGAATAACTGAGCGCTGA 67
 DB 8 GAGCGCTAACGCTCTTCTGTCTCCCGCGGTGTGTCATGACGGTGAATAACTGAGCGCTGA 67
 QY 68 AGGCGACCTCTCACTTACTCCAGGATGAGGGCGCTGGTGGCAATTCATCGCTTCATGA 127
 DB 68 AGGCGACCTCTCACTTACTCCAGGATGAGGGCGCTGGTGGCAATTCATCGCTTCATGA 127
 QY 128 AGCAGAGGAGGATGAGGCTGTAACGACTTTTATTCAGAGATTTGCAATTAACCTTCATGCAT 187
 DB 128 AGCAGAGGAGGATGAGGCTGTAACGACTTTTATTCAGAGATTTGCAATTAACCTTCATGCAT 187
 QY 188 GCAACACCCCTGAAGCTTCACTCCCTTGAAGACTTCCCAACCTCAGAGCGCTGAGCTTA 247
 DB 188 GCAACACCCCTGAAGCTTCACTCCCTTGAAGACTTCCCAACCTCAGAGCGCTGAGCTTA 247

Qy	248	TGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCTCAGCAAAATCAACCTTTGGGCCCGTCGT	307
Db	248	TGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCTCAGCAAAATCAACCTTTGGGCCCGTCGT	307
Qy	308	CCAATCCTCATGCTAAACCACTCTGACATTTTCACTTCTTTGAAAGTCATCGGAAAGGCGCAGTT	367
Db	308	CCAATCCTCATGCTAAACCACTCTGACATTTTCACTTCTTTGAAAGTCATCGGAAAGGCGCAGTT	367
Qy	368	TTGGAAGAGTTCTTCTTAGCAAGACACAAGGCGAGAAGAGTGTTCATATGCAGTCAAAAGTTT	427
Db	368	TTGGAAGAGTTCTTCTTAGCAAGACACAAGGCGAGAAGAGTGTTCATATGCAGTCAAAAGTTT	427
Qy	428	TACAGAAGAAGCAATCCTCGAANAAGAAAGAGGAAGCATATATATGTCGAGCGCGGAATG	487
Db	428	TACAGAAGAAGCAATCCTCGAANAAGAAAGAGGAAGCATATATATGTCGAGCGCGGAATG	487
Qy	488	TTCTGTTGAAGAATGTGAAGACCCCTTTCCTGTGTGGGCTTTCACCTCTCTTTCCAGACTG	547
Db	488	TTCTGTTGAAGAATGTGAAGACCCCTTTCCTGTGTGGGCTTTCACCTCTCTTTCCAGACTG	547
Qy	548	CTGACAAATGTACTTTTGTCTCTAGACATACATTAATGCTGGAGAGTTGTTCTTACCATCTCC	607
Db	548	CTGACAAATGTACTTTTGTCTCTAGACATACATTAATGCTGGAGAGTTGTTCTTACCATCTCC	607
Qy	608	AGAGGGAACGCTGCTTCTCGAACACCGGGCTCTTTCTATATGCTGTGTAANAATAGCCAGTG	667
Db	608	AGAGGGAACGCTGCTTCTCGAACACCGGGCTCTTTCTATATGCTGTGTAANAATAGCCAGTG	667
Qy	668	CTTTGGGCTACCTGCTATCTCACTGAACATCGTTTATAGAGACTTAAAAACACAGAGATATTT	727
Db	668	CTTTGGGCTACCTGCTATCTCACTGAACATCGTTTATAGAGACTTAAAAACACAGAGATATTT	727
Qy	728	TGCTAGATTACAGGGACACATTTGTCCTTACTGATTTTCGGACTCTGCAAGGAGAACATTTG	787
Db	728	TGCTAGATTACAGGGACACATTTGTCCTTACTGACTTCGGACTCTGCAAGGAGAACATTTG	787
Qy	788	AACACACACGACAAACATCCACCTCTCTGGCAGCGCGGAGTATCTCGCACCTTGAGTGC	847
Db	788	AACACACACGACAAACATCCACCTCTCTGGCAGCGCGGAGTATCTCGCACCTTGAGTGC	847
Qy	848	TTCATAGCAGCCCTTATCAGAGGACTGTGAGCTGCTGGTGCTGGGAGCTGCTCTGTATG	907
Db	848	TTCATAGCAGCCCTTATCAGAGGACTGTGAGCTGCTGGTGCTGGGAGCTGCTCTGTATG	907
Qy	908	AGATGCTGTATGGCTCGCCGCTTTTATAGCCGAAACACACAGCTGAAATGTACGACAACA	967
Db	908	AGATGCTGTATGGCTCGCCGCTTTTATAGCCGAAACACACAGCTGAAATGTACGACAACA	967
Qy	968	TTCTGAAACAGCCCTCTCCAGCTGAAACCAATATTACAATTCGCGAAGACACCTCTCTGG	1027
Db	968	TTCTGAAACAGCCCTCTCCAGCTGAAACCAATATTACAATTCGCGAAGACACCTCTCTGG	1027
Qy	1028	AGGGCTTCTTCAGAGGACAGGACAAAGCGGCTCGGGGCGCAAGGATGACTTTCATGAGA	1087
Db	1028	AGGGCTTCTTCAGAGGACAGGACAAAGCGGCTCGGGGCGCAAGGATGACTTTCATGAGA	1087
Qy	1088	TTAAGAGCTATGCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTA	1147
Db	1088	TTAAGAGCTATGCTTCTCTCTCTTAATTAACCTGGGATGATCTCATTAAATAAGAGATTA	1147
Qy	1148	CTCCCCCTTTTAAACCCAAATGTGAGTGGGCGCCCAACGAGCTACGSCACTTTGACCCCGAGT	1207
Db	1148	CTCCCCCTTTTAAACCCAAATGTGAGTGGGCGCCCAACGAGCTACGSCACTTTGACCCCGAGT	1207
Qy	1208	TTACCGAAGAGCCTGTCCCCAACTCCATTGGCAGTCCCCCTGACAGGCTCTCTGTACAG	1267
Db	1208	TTACCGAAGAGCCTGTCCCCAACTCCATTGGCAGTCCCCCTGACAGGCTCTCTGTACAG	1267
Qy	1268	CCAGCGTCAAGGAAGCTGCCGAGGCTTTCCTTAGGCTTTTCTCTATATGCGCCTCCACAGCACT	1327
Db	1268	CCAGCGTCAAGGAAGCTGCCGAGGCTTTCCTTAGGCTTTTCTCTATATGCGCCTCCACAGCACT	1327
Qy	1328	CTTTCCTCTGAACCCCTGTTAAGGCGTTGGTTTTTAAAGGATTTTATGTGTGTGTTCCGGAATG	1387

AK098509	AK098509	2364 bp	mrna	linear	PRI 15-JUL-2002
LOCUS	Homo sapiens cDNA FLJ25643 fis, clone STM07148, highly similar to				
DEFINITION	SERINE/THREONINE-PROTEIN KINASE SGK (EC 2.7.1.1.)				
ACCESSION	AK098509				
VERSION	AK098509.1	GI:21758535			
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.				
TITLE	NEDO human cDNA sequencing project				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 2364)				
AUTHORS	Sugano, S. and Suzuki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: cdnalgims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)				
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.				
FEATURES	Location/Qualifiers				
Source	1. .2364				
	/organism="Homo sapiens"				
	/mol_type="mrna"				
	/db_xref="taxon:9606"				
	/clone="STM07148"				
	/tissue_type="stomach mucosa"				
	/clone_lib="STM"				
	/note="cloning vector: pWE18sFL3"				
BASE COUNT	615 a 524 c 517 g 708 t				
ORIGIN					
Query Match	97.5%; Score 2311.6; DB 9; Length 2364;				
Best Local Similarity	99.7%; Pred. No. 0;				
Matches 2336; Conservative	0; Mismatches 4; Indels 2; Gaps 2;				
QY	8 GAGCGCTAAGCTTCTCTCTCCCGCGGTGGTGATGACGGTGAAGAACTGAGGCTGCTA 67				
DB					
DB	23 GAGCGCTAAGCTTCTCTCTCCCGCGGTGGTGATGACGGTGAAGAACTGAGGCTGCTA 82				
QY	68 AGGGACCTCTACTTACTCCAGATGAGGGGATGGGCAATCTCTATCGCTTTCATGA 127				
DB					
DB	83 AGGGACCTCTACTTACTCCAGATGAGGGGATGGGCAATCTCTATCGCTTTCATGA 142				
QY	128 AGCAGAGGAGGTGGTCTGAACGACTTTATTTCAGAGATTTGCCAATAACTCTATGCAT 187				
DB					
DB	143 AGCAGAGGAGGTGGTCTGAACGACTTTATTTCAGAGATTTGCCAATAACTCTATGCAT 202				
QY	188 GCNAACACCTCTGAAGTTCACTTCATCTTGAAGATCTCCCAACTCAGAGCCCTGAGCTTA 247				
DB					
DB	203 GCNAACACCTCTGAAGTTCACTTCATCTTGAAGATCTCCCAACTCAGAGCCCTGAGCTTA 262				
QY	248 TGAATCCACCTCTCTCTCCACCAAGTCTCTCTCAGCAATCACTTGGCCCGCTGCT 307				
DB					
DB	263 TGAATCCACCTCTCTCTCTCCACCAAGTCTCTCTCAGCAATCACTTGGCCCGCTGCT 322				
QY	308 CCAATCCTCATGCTAAACCATCTGACTTTTCACTTCTTGAAAGTGATCGGAAGGGCAGTT 367				

Db	323 CCAATCCTCATGCTAAACCATCTGACTTTTCACTTCTTGAAAGTGATCGGAAGGGCAGTT 382				
QY	368 TTGGAAGCTTCTTCTAGCAACACACAGGAGAGAAAGTGTCTTCTATCGACTCAAGTTT 427				
Db	383 TTGGAAGGTTCTTCTAGCAACACACAGGAGAGAAAGTGTCTTCTATCGACTCAAGTTT 442				
QY	428 TACAGAAGAAGCAATCTGAAAAAGAGAGAGAGCATATTATGTCGAGAGCAAGATG 487				
Db	443 TACAGAAGAAGCAATCTGAAAAAGAGAGAGAGCATATTATGTCGAGAGCAAGATG 502				
QY	488 TTCTCTTGAAGAATCTGAAGACCTTCTCTGGTGGGCTTTCACCTCTCTTTCAGAGT 547				
Db	503 TTCTCTTGAAGAATCTGAAGACCTTCTCTGGTGGGCTTTCACCTCTCTTTCAGAGT 562				
QY	548 CTGCAAAATGTACTTCTCTAGACTACATTAAGTGGTGGAGAGTGTCTTACCAATCT 607				
Db	563 CTGCAAAATGTACTTCTCTAGACTACATTAAGTGGTGGAGAGTGTCTTACCAATCT 622				
QY	608 AGAGGGAAGCTGTCTTCTGGAACACAGGGCTCGTTTCTATGCTCTGAAATAGTCAAG 667				
Db	623 AGAGGGAAGCTGTCTTCTGGAACACAGGGCTCGTTTCTATGCTCTGAAATAGTCAAG 682				
QY	668 CCTTGGGCTACCTGCAATTCACATGACATCGTTTATAGAGACTTAAATCAGAGAAATTT 727				
Db	683 CCTTGGGCTACCTGCAATTCACATGACATCGTTTATAGAGACTTAAATCAGAGAAATTT 742				
QY	728 TGCTAGATTCACAGGACACATGTCTTACTGATTTCCGACTCTGCAAGAGTAATATG 787				
Db	743 TGCTAGATTCACAGGACACATGTCTTACTGACTTCGGACTCTGCAAGAGAAATATG 802				
QY	788 AACACAACAGCACATCCACCTTCTGTGGACCCGCGAGTATCTCGACCTGAGGTGC 847				
Db	803 AACACAACAGCACATCCACCTTCTGTGGACCCGCGAGTATCTCGACCTGAGGTGC 862				
QY	848 TTCATAAGCAGCCTTATGACAGACTGTGGACTGTGGTGCCTGGGAGCTGTCTTGTATG 907				
Db	863 TTCATAAGCAGCCTTATGACAGACTGTGGACTGTGGTGCCTGGGAGCTGTCTTGTATG 922				
QY	908 AGATGCTGTATGGCTTGGCGCTTTTATAGCGAAACACACACTGAAATGACAGTAATTA 967				
Db	923 AGATGCTGTATGGCTTGGCGCTTTTATAGCGAAACACACACTGAAATGACAGTAATTA 982				
QY	968 TTCTGAACAAGCTCTCCAGCTGAAACCAATATTACAAA TTCCGCAAGACAAATCTCTG 1027				
Db	983 TTCTGAACAAGCTCTCCAGCTGAAACCAATATTACAAA TTCCGCAAGACAACTCTCTG 1042				
QY	1028 AGGGCTCTCTCAGAGACAGCAACAGCGCTCGGGCCCAAGGATGACTTTCATGGAGA 1087				
Db	1043 AGGGCTCTCTCAGAGACAGCAACAGCGCTCGGGCCCAAGGATGACTTTCATGGAGA 1102				
QY	1088 TTAAGAGTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1147				
Db	1103 TTAAGAGTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1162				
QY	1148 CTCCTCCCTTTTAAACCAATGTAGTGGCCCAACAGCTAGCGACTTTGACCCCGAGT 1207				
Db	1163 CTCCTCCCTTTTAAACCAATGTAGTGGCCCAACAGCTAGCGACTTTGACCCCGAGT 1222				
QY	1208 TTACCGAAGAGCTGTCTCCCAACTCCATTGGCAAGTCCCTCAGACAGCTCTCTGTCACAG 1267				
Db	1223 TTACCGAAGAGCTGTCTCCCAACTCCATTGGCAAGTCCCTCAGACAGCTCTCTGTCACAG 1282				
QY	1268 CCAGCGTCAAGAGCTGCGGAGGCTTTCTTAGGCTTTTCTTATGCGGCTCTCCACGAGCT 1327				
Db	1283 CCAGCGTCAAGAGCTGCGGAGGCTTTCTTAGGCTTTTCTTATGCGGCTCTCCACGAGCT 1342				
QY	1328 CTTTCTCTGACACCTGTGTAGGCTGTGGTTTAAAGGATTTTATGTGTGTCTTCCAAATGT 1387				
Db	1343 CTTTCTCTGACACCTGTGTAGGCTGTGGTTTAAAGGATTTTATGTGTGTCTTCCAAATGT 1402				
QY	1388 TTTAGTTAGCCTTTTGTGTGAGGCTGCGGAGCTGACAGGACATCTTACAAAGAGAAATTTGCAC 1447				

QY	874	GTGACTGGTGGTGGAGCTGTCTTGTATGAGATGCTGTATGGCTGCCCTTTT	933	QY	1951	AATGCTCTGCAATGACTGTAATTCAGATGCTTAAAGAAAGCATTTGCTGTACAAATTT	2010
DB	841	GTGACTGGTGGTGGAGCTGTCTTGTATGAGATGCTGTATGGCTGCCCTTTT	900	DB	1921	AATGCTCTGCAATGACTGTAATTCAGATGCTTAAAGAAAGCATTTGCTGTACAAATTT	1980
QY	934	TATAGCGGAACACAGCTGAATGTAGACAACATTCGAACAGCCTCTCCAGCTGAAA	993	QY	2011	TCATATTTTGAAGAGGTTTTTATGACCAATGCCAGTTGTGTCAGTCAGAGCTGTGGT	2070
DB	901	TATAGCGGAACACAGCTGAATGTAGACAACATTCGAACAGCCTCTCCAGCTGAAA	960	DB	1981	TCATATTTTGAAGAGGTTTTTATGACCAATGCCAGTTGTGTCAGTCAGAGCTGTGGT	2040
QY	994	CCAAATATTACAAATCCGGAAGACACTCTCTGAGGGCTCTGCGAGAGGACAGACGA	1053	QY	2071	GTATTTTCATGCTTTAAATGTCACCTGTAAATGGGCATTTATTTATCTTTTTTTTTC	2130
DB	961	CCAAATATTACAAATCCGGAAGACACTCTCTGAGGGCTCTGCGAGAGGACAGACGA	1020	DB	2041	GTATTTTCATGCTTTAAATGTCACCTGTAAATGGGCATTTATTTATCTTTTTTTTTC	2100
QY	1054	AAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATGCTTCTTCTCCTTA	1113	QY	2131	ATTCCTGATATGCTATGTATTAAGAAAGCTCTGTACATTTGGCTTAAACATTTAGT	2190
DB	1021	AAGCGCTCGGGCCCAAGGATGACTTCATGAGATTAAAGAGTCATGCTTCTTCTCCTTA	1080	DB	2101	ATTCCTGATATGCTATGTATTAAGAAAGCTCTGTACATTTGGCTTAAACATTTAGT	2160
QY	1114	ATTAACCTGGGATGATCTCAATTAAGAAGATTACTCCCTTTTAAACCAATTTAGT	1173	QY	2191	ATATTTAACTTACAGGCTTATTTGTATTAACCAACCATTTTAAATGTACTGTAATTA	2250
DB	1081	ATTAACCTGGGATGATCTCAATTAAGAAGATTACTCCCTTTTAAACCAATTTAGT	1140	DB	2161	ATATTTAACTTACAGGCTTATTTGTATTAACCAACCATTTTAAATGTACTGTAATTA	2220
QY	1174	GGGCCCAACAGCTACGGCTTTTACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCC	1233	QY	2251	CATGCTTATTAATAGTACCAATCTCTTCCCTCATCCCATCACACACTTTTTTTTGTGTGA	2310
DB	1141	GGGCCCAACAGCTACGGCTTTTACCCCGAGTTTACCGAAGAGCCTGTCCCAACTCC	1200	DB	2221	CATGCTTATTAATAGTACCAATCTCTTCCCTCATCCCATCACACACTTTTTTTTGTGTGA	2280
QY	1234	ATTGGCAAGTCCCTGACAGGCTCTGTCACGCCAGGCTCAAGGAAGCTGCCGAGCT	1293	QY	2311	TAACTGATTTGGTTTGGCAATAAAACCTTG	2341
DB	1201	ATTGGCAAGTCCCTGACAGGCTCTGTCACGCCAGGCTCAAGGAAGCTGCCGAGCT	1260	DB	2281	TAACTGATTTGGTTTGGCAATAAAACCTTG	2311
QY	1294	TTCTAGGCTTTTCTATGGCCCTCCACGAGCTTTTCTCTCAACCCCTTTAGGCTT	1353	RESULT 11			
DB	1261	TTCTAGGCTTTTCTATGGCCCTCCACGAGCTTTTCTCTCAACCCCTTTAGGCTT	1320	LOCUS	AR270209	2311 bp	DNA
QY	1354	GGTTTTAAAGATTTTATGCTGTGTTCCGAATGTTTGTAGTGGTGGAGCCG	1413	DEFINITION	Sequence 772 from patent US 6500938.		linear
DB	1321	GGTTTTAAAGATTTTATGCTGTGTTCCGAATGTTTGTAGTGGTGGAGCCG	1380	ACCESSION	AR270209		
QY	1414	CAGCTGACAGACATCTACAAGAGATTTGACATCTCTGGAGCTTACCAATCTATT	1473	VERSION	AR270209.1	GI:29701443	
DB	1381	CAGCTGACAGACATCTACAAGAGATTTGACATCTCTGGAGCTTACCAATCTATT	1440	KEYWORDS	unknown.		
QY	1474	GCACACTGTCGTGGGAA-TTTTGTGAAGACATTTCTCTCAGTCAAGCTCATGAGTT	1532	ORGANISM	Unknown.		
DB	1441	GCACACTGTCGTGGGAA-TTTTGTGAAGACATTTCTCTCAGTCAAGCTCATGAGTT	1500	REFERENCE	1 (bases 1 to 2311)		
QY	1533	TTCAATTTTATCTCTTCCATGCTGCTGTAATGCTGTAACAGAGGCTTAGAGTGGCG	1592	AUTHORS	An-Young, J. and Seilhamer, J.J.		
DB	1501	TTCAATTTTATCTCTTCCATGCTGCTGTAATGCTGTAACAGAGGCTTAGAGTGGCG	1560	TITLE	Composition for the detection of signaling pathway gene expression		
QY	1593	CTTAGACGAGGAGGAGGTTTCGTTAGAAAGCGGAC-CTGTTCTAAAAAGGCTCCCTGC	1651	JOURNAL	Patent: US 6500938-A 772 31-DEC-2002;		
DB	1561	CTTAGACGAGGAGGAGGTTTCGTTAGAAAGCGGACCTGTTCTAAAAAGGCTCCCTGC	1620	FEATURES	Location/Qualifiers		
QY	1652	AGATCTGCTGGGCTGTGATGACGAATATTATGAATGTGCCCTTTCTGAAG-AGATTGT	1710	source	1. 2311		
DB	1621	AGATCTGCTGGGCTGTGATGACGAATATTATGAATGTGCCCTTTCTGAAG-AGATTGT	1680	BASE COUNT	604 a 508 c 506 g 692 t	1 others	
QY	1711	GTATGCTCCAAAGCTTTCTATGCGAGCTGTTTCAAGTCTTTTATTTCCCTTGTGGATAT	1770	ORIGIN			
DB	1681	GTATGCTCCAAAGCTTTCTATGCGAGTGTTCAGTCTTTTATTTCCCTTGTGGATAT	1740	Query Match	95.5%; Score 2262.8; DB 6; Length 2311;		
QY	1771	GCTGTGAAACCTGCTGTGATGCTGATGCTGATCAGATGATGATTTGTTATTAAGC	1830	Best local Similarity	99.5%; Pred. No. 0;		
DB	1741	GCTGTGAAACCTGCTGTGATGCTGATGCTGATCAGATGATGATTTGTTATTAAGC	1800	Matches 2300; Conservative	0; Mismatches 8; Indels 3; Gaps 3;		
QY	1831	ATCAATGTGACACTTGCAGGACACTACAACGTGGGACATTTGTTTCTTCCATATTTG	1890	QY	34	CGGCTGTGATGACGGTGAAACTGAGGCTGCTAAGGGCACCTCTACTTACTCCAGATG	93
DB	1801	ATCAATGTGACACTTGCAGGACACTACAACGTGGGACATTTGTTTCTTCCATATTTG	1860	DB	1	CGGCTGTGATGACGGTGAAACTGAGGCTGCTAAGGGCACCTCTACTTACTCCAGATG	60
QY	1891	GAGATAAATTTATGTGTAGACTTTTGTGAAGTACGGTTTAACTAAATTTATTTGA	1950	QY	94	AGGGGCATGCTGGCAATTTCTATCGCTTTTCATGAAGCAGAGGAGGATGGCTCTGAACGAC	153
DB	1861	GAGATAAATTTATGTGTAGACTTTTGTGAAGTACGGTTTAACTAAATTTATTTGA	1920	DB	61	AGGGGCATGCTGGCAATTTCTATCGCTTTTCATGAAGCAGAGGAGGATGGCTCTGAACGAC	120

QY	334	TTTCACTTCTTGAAGTGA	TCGAAAGGGCAGTTT	TGGAAAGTCTTCTT	AGCAAGACAC	393
DB	301	TTTCACTTCTTGAAGTGA	TCGAAAGGGCAGTTT	TGGAAAGTCTTCTT	AGCAAGACAC	360
QY	394	AAGGCAGAGAAGTGTCT	TATGCAGTCAAAAGTTT	TACAGAAGAACAAT	TCCTGAAAAG	453
DB	361	AAGGCAGAGAAGTGTCT	TATGCAGTCAAAAGTTT	TACAGAAGAACAAT	TCCTGAAAAG	420
QY	454	AAGAGAGGAGAGCATAT	TATTCGGAGCGGATGCT	TGTTGAAGAAATGTGA	AGCACCCCT	513
DB	421	AAGAGAGGAGAGCATAT	TATTCGGAGCGGATGCT	TGTTGAAGAAATGTGA	AGCACCCCT	480
QY	514	TTCCCTGGTGGCCCTTC	ACTTCTCTTCCAGACTGC	TGACAAATGTGTACT	TTTGCTCCAGAC	573
DB	481	TTCCCTGGTGGCCCTTC	ACTTCTCTTCCAGACTGC	TGACAAATGTGTACT	TTTGCTCCAGAC	540
QY	574	TACATTAAATCGTGGAG	AGTTGTTCTACCATCTCC	AGAGGAAAGCGTCTCT	CTTGGAAACCA	633
DB	541	TACATTAAATCGTGGAG	AGTTGTTCTACCATCTCC	AGAGGAAAGCGTCTCT	CTTGGAAACCA	600
QY	634	CGGGCTCGTTTCTATGC	TGTAATAGGCAGTGCCT	TGGGCTACCTGCAAT	TTCACCTGAAC	693
DB	601	CGGGCTCGTTTCTATGC	TGTAATAGGCAGTGCCT	TGGGCTACCTGCAAT	TTCACCTGAAC	660
QY	694	ATCGTTTATAGAGACTTA	AAACCCAGAGAAATATT	TGCTAGATTACAGG	GACACATGTGC	753
DB	661	ATCGTTTATAGAGACTTA	AAACCCAGAGAAATATT	TGCTAGATTACAGG	GACACATGTGC	720
QY	754	CTTACTGATTTCGGACTC	TCGAAGGAGAACATTGA	ACAACAACAGACACAT	CCACCTTC	813
DB	721	CTTACTGATTTCGGACTC	TCGAAGGAGAACATTGA	ACAACAACAGACACAT	CCACCTTC	780
QY	814	TGTGGCACGCGGAGTAT	CTCGCACTTGAGTGCCT	CATTAAGCAGCCTTAT	GACAGGACT	873
DB	781	TGTGGCACGCGGAGTAT	CTCGCACTTGAGTGCCT	CATTAAGCAGCCTTAT	GACAGGACT	840
QY	874	GTGACTGTGTGTCCTCG	GAGCTGTCTGTATGAGAT	GTCTGTATGBCCTG	CCGCGCTTTT	933
DB	841	GTGACTGTGTGTCCTCG	GAGCTGTCTGTATGAGAT	GTCTGTATGBCCTG	CCGCGCTTTT	900
QY	934	TATAGCCGAACAACAGC	TGAAATGTACGACAACAT	TCTGAAACAAGCCTC	TCCAGCTGAAA	993
DB	901	TATAGCCGAACAACAGC	TGAAATGTACGACAACAT	TCTGAAACAAGCCTC	TCCAGCTGAAA	960
QY	994	CCAATATTACAAATTCG	GACACACCTCTCGAGG	CGCCTCTGCAAGAG	CACAGGACA	1053
DB	961	CCAATATTACAAATTCG	GACACACCTCTCGAGG	CGCCTCTGCAAGAG	CACAGGACA	1020
QY	1054	AAGCGCTCGGGCCAGAG	TGACTTCATGAGATTAA	GAGTATGCTTCTCTC	CTCTTA	1113
DB	1021	AAGCGCTCGGGCCAGAG	TGACTTCATGAGATTAA	GAGTATGCTTCTCTC	CTCTCTTA	1080
QY	1114	ATTAACTGGATGATCTCA	TATTAAGAATACTCC	CCCTTTTACCCCAAT	GTGAGT	1173
DB	1081	ATTAACTGGATGATCTCA	TATTAAGAATACTCT	CCCCCTTTTACCCCAAT	GTGAGT	1140
QY	1174	GGGCCCAACAGCTACG	GCACCTTTGACCCGAG	TTTACCGAAGAGC	CTGTCCCAACTCC	1233
DB	1141	GGGCCCAACAGCTACG	GCACCTTTGACCCGAG	TTTACCGAAGAGC	CTGTCCCAACTCC	1200
QY	1234	ATTGGCAAGTCCCTGAC	AGGTCCTGTCACAGC	ACGCTCAAGGAGCT	GTCCGAGGCT	1293
DB	1201	ATTGGCAAGTCCCTGAC	AGGTCCTGTCACAGC	ACGCTCAAGGAGCT	GTCCGAGGCT	1360
QY	1294	TTCCTAGGCTTTTCC	TATGCGCTCCCA	CGSACTCTTCTCTG	AACCCCTGTTAGG	1353
DB	1261	TTCCTAGGCTTTTCC	TATGCGCTCCCA	CGSACTCTTCTCTG	AACCCCTGTTAGG	1320
QY	1354	GGTTTAAAGGATTTAT	GTGTGTTCCGAAATG	TTTAGTTAGCCTTT	TGGTGGAGCCG	1413
DB	1321	GGTTTAAAGGATTTAT	GTGTGTTCCGAAATG	TTTAGTTAGCCTTT	TGGTGGAGCCG	1380
QY	1414	CAGCTGACAGGACATC	TACAAGAGAATTTG	CAATCTCTGGAAGC	TTAGCAATCTTATT	1473

[illegible]

RESULT 12
BD021928

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

2311 bp DNA linear PAT 27-AUG-2002

QY	1771	GCTGTGTGAACCGCTGCTGTGAGTGTGCTGATGCTGATCAGAGATGATTTCTTTATAAGC	1830	118	GCTTTCATGAAGCAGAGAGAGATGGTCTGAACGACTTTTATTCAGAGATTTGCAATTAAC	177	
DB	1741	CTGTGTGAACCGCTGCTGTGAGTGTGCTGATGCTGATCAGAGATGATTTCTTTATAAGC	1800	959	GCITTCATGAAGCAGAGAGAGATGGTCTGAACGACTTTTATTCAGAGATTTGCAATTAAC	1018	
QY	1831	ATCAATGTGACACTTCGAGACACTACACGCTGGGACATTTGTTTCTTCCATATTG	1890	178	TCCTATGATGATCAACACACCCCTGAAGTTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAG	237	
DB	1801	ATCAATGTGACACTTCGAGACACTACACGCTGGGACATTTGTTTCTTCCATATTG	1860	1019	TCCTATGATGATCAACACACCCCTGAAGTTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAG	1078	
QY	1891	GAAGATAAATTTATGTGTAGACTTTTATGTAAGTACGGTTATTAACCTAAATTTATTGA	1950	238	CCTGAGCTTATCAATGGTCAACCCCTTCTCTCCACCAAGTCTTCTTCAGCAAAATCAATCTT	297	
DB	1861	GAAGATAAATTTATGTGTAGACTTTTATGTAAGTACGGTTATTAACCTAAATTTATTGA	1920	1079	CCTGAGCTTATCAATGGTCAACCCCTTCTCTCCACCAAGTCTTCTTCAGCAAAATCAATCTT	1148	
QY	1951	ATGGCTTCGCAATGACTGCTATTCAGATGCTTAAGAGCAATTCCTGCTCAACATATT	2010	298	GGCCGCTGCTCAATGCTCAATGCTTAAGAGCAATTCCTGCTCAACATTTCTTCAAGTCTCAAGCA	357	
DB	1921	ATGGCTTCGCAATGACTGCTATTCAGATGCTTAAGAGCAATTCCTGCTCAACATATT	1980	1139	GGCCGCTGCTCAATGCTCAATGCTTAAGAGCAATTCCTGCTCAACATTTCTTCAAGTCTCAAGCA	1198	
QY	2011	TCATTTTATGAAGGGTTTTTATGACCAATGCCCGAGTTGTCTCAGAGCCGTTGGT	2070	358	AAGGCACTTTTGAAGAGTTCTTCTAGCAAGACACAGGAGGAGAGAGTGTCTTATGCA	417	
DB	1981	TCATTTTATGAAGGGTTTTTATGACCAATGCCCGAGTTGTCTCAGAGCCGTTGGT	2040	1199	AAGGCACTTTTGAAGAGTTCTTCTAGCAAGACACAGGAGGAGAGAGTGTCTTATGCA	1258	
QY	2071	GTTTTTCATTTTAAATGCTACCTGTAAATGGGCATTTATTTTCTTCTTCTTCTTCT	2130	418	GTCAAGTTTTCAG	477	
DB	2041	GTTTTTCATTTTAAATGCTACCTGTAAATGGGCATTTATTTTCTTCTTCTTCTTCT	2100	1259	GTCAAGTTTTCAG	1318	
QY	2131	ATTCCTGATATTTATGTATTTATTAAGAGAGCTGTGTACATTTGGGTTTATACACTAGT	2190	478	GAGCGGAATGTTCTGTTGAAGAGTTGTGAAGCACCCTTTCTCTGGTGGGCTTCTACATCTCT	537	
DB	2101	ATTCCTGATATTTATGTATTTATTAAGAGAGCTGTGTACATTTGGGTTTATACACTAGT	2160	1319	GAGCGGAATGTTCTGTTGAAGAGTTGTGAAGCACCCTTTCTCTGGTGGGCTTCTACATCTCT	1378	
QY	2191	ATATTTAACTTACAGGCTTATTTGTAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2250	538	TTCCAGAGCTGCTGACAAATTTGACTTTTCTCTAGACACTACATTAATGTTGGAGAGAGTCTTC	597	
DB	2161	ATATTTAACTTACAGGCTTATTTGTAATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2220	1379	TTCCAGAGCTGCTGACAAATTTGACTTTTCTCTAGACACTACATTAATGTTGGAGAGAGTCTTC	1438	
QY	2251	CATGGTTATATAGTACATTCCTTCTCCTCATCCCATCACACAACTTTTCTGTGTGTA	2310	598	TACCATCTCCAGAGGAG	657	
DB	2221	CATGGTTATATAGTACATTCCTTCTCCTCATCCCATCACACAACTTTTCTGTGTGTA	2280	1439	TACCATCTCCAGAGGAG	1498	
QY	2311	TAACTGATTTTGGTTTGGCAATTAACCTTTG	2341	658	ATAGCCAGTGGCTTGGGCTTACCTGCAATTCAGTGAACATCTGTTTATAGAGAGAGTAAAGCA	717	
DB	2281	TAACTGATTTTGGTTTGGCAATTAACCTTTG	2311	1499	ATAGCCAGTGGCTTGGGCTTACCTGCAATTCAGTGAACATCTGTTTATAGAGAGAGTAAAGCA	1558	
RESULT 13				QY	718	GAGAATATTTCTGCTAGATTTCAGAGGACACATTTGCTTACTGATTTTCGAGACTCTGGAAG	777
LOCUS	AX713538	3196 bp	DNA	linear	PAT 15-APR-2003		
DEFINITION	Sequence 222 from Patent EP1293569.			DB	1559	GAGAATATTTCTGCTAGATTTCAGAGGACACATTTGCTTACTGATTTTCGAGACTCTGGAAG	1618
ACCESSION	AX713538			QY	778	GAGAATATTTCTGCTAGATTTCAGAGGACACATTTGCTTACTGATTTTCGAGACTCTGGAAG	837
VERSION	AX713538.1	GI:29888396		DB	1619	GAGAATATTTCTGCTAGATTTCAGAGGACACATTTGCTTACTGATTTTCGAGACTCTGGAAG	1678
KEYWORDS				QY	838	CCTGAGGCTGCTTCAATAGCAGCTTATGACAGGAGCTGAGCTGGTGGTGGTGGGAGCT	897
SOURCE	Homo sapiens (human)			DB	1679	CCTGAGGCTGCTTCAATAGCAGCTTATGACAGGAGCTGAGCTGGTGGTGGTGGGAGCT	1738
ORGANISM	Homo sapiens			QY	898	GTCTGTATGATGATGCTGTATGGCTGCTGGCTCTTTTATAGCCGAAACACAGCTTGAATG	957
REFERENCE	Patent: EP 1293569-A 222 19-MAR-2003;			DB	1739	GTCTGTATGATGATGCTGTATGGCTGCTGGCTCTTTTATAGCCGAAACACAGCTTGAATG	1798
AUTHORS	Helix Research Institute (JP) ; Research Association for Biotechnology (JP)			QY	958	TAGCACACATTTCTGACAGAGCTCTCCAGCTGAAACCAATATATACAAATTTCCGCAAGA	1017
TITLE	Full-length cDNAs			DB	1799	TAGCACACATTTCTGACAGAGCTCTCCAGCTGAAACCAATATATACAAATTTCCGCAAGA	1858
JOURNAL	Location/Qualifiers			QY	1018	CACCTCTCGAGAGGCTCTCTGAG	1077
FEATURES	1. 3196			DB	1859	CACCTCTCGAGAGGCTCTCTGAG	1918
source	/organism="Homo sapiens"			QY	1078	TTATGAGAGATTAAAGAGTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1137
	/mol_type="genomic DNA"			DB	1919	TTATGAGAGATTAAAGAGTATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1978
	/db_xref="taxon:9606"			QY	1138	AAGAAGATTACTCCCTCTTTTAAACCAATATGAGTGGGCCCCCAACAGCTTACGGCACTTT	1197
	817 a 767 c 723 g 889 t			DB	1979	AAGAAGATTACTCCCTCTTTTAAACCAATATGAGTGGGCCCCCAACAGCTTACGGCACTTT	2038
BASE COUNT	817 a 767 c 723 g 889 t						
ORIGIN							
Query Match	92.9%	Score 2201;	DB 6;	Length 3196;			
Bert Local Similarity	99.7%	Pred. No. 0;					
Matches 2226;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 2;			

[illegible]

QY 1993 ATTCTGCTACAAATATTTCTATTTTAGAAAGGGTTTTATGACCAATGCCCCAGTTG 2052
DB |||||||
QY 1895 ATTCTGCTACAAATATTTCTATTTTAGAAAGGGTTTTATGACCAATGCCCCAGTTG 1954
DB |||||||
QY 2053 TCAGTCACAGCCGTTGGTGTTCATTTTCAATGTTTAAATGTCACCTGTAAATGGGCATTAT 2112
DB |||||||
QY 1955 TCAGTCACAGCCGTTGGTGTTCATTTTCAATGTTTAAATGTCACCTGTAAATGGGCATTAT 2014
DB |||||||
QY 2113 TTATGTTTTTTTTTTCGATTCCTGATAATTTGATGATGATAAAGACGTCCTGTACA 2172
DB |||||||
QY 2015 TTATGTTTTTTTTTTCGATTCCTGATAATTTGATGATGATAAAGACGTCCTGTACA 2074
DB |||||||
QY 2173 TTGGGTTATACACTAGTATATTTAAACTTACAGGCTTATTTGTAATGTAACACCACTT 2232
DB |||||||
QY 2075 TTGGGTTATACACTAGTATATTTAACTTACAGGCTTATTTGTAATGTAACACCACTT 2134
DB |||||||
QY 2233 TTAATGCTACTGTAATTAACATGGTTATTAATACGTAACATCCCTCCCTCATCCCATCACAC 2292
DB |||||||
QY 2135 TTAATGCTACTGTAATTAACATGGTTATTAATACGTAACATCCCTCCCTCATCCCATCACAC 2194
DB |||||||
QY 2293 AACTTTTTTTGTGTGATAAAGCTGATTTTGGTTGCAATAAAGACCTTGAAAAATAAAA 2352
DB |||||||
QY 2195 AACTTTTTTTGTGTGATAAAGCTGATTTTGGTTGCAATAAAGACCTTGAAAAATAAAA 2254
DB |||||||
QY 2353 AAAAAAAAAAAAAA 2368
DB |||||||
QY 2255 AAAAAAAAAAAAAA 2270
DB |||||||

Search completed: August 11, 2003, 04:53:03
Job time : 8496 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 01:25:43 ; Search time 616 Seconds
(without alignments)
10385.825 Million cell updates/sec

Title: US-10-000-039a-1
Perfect score: 2370
Sequence: 1 CACGAGGAGCCCTACGTC.....AAAAAAAAAAAAAAAAAAAA 2370

Scoring table: IDENTITY_NUC
Gapop 10.0 , Capext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2370	100.0	2370	19 AAV48311	Human cell-volume
2	2370	100.0	2370	24 ABR34749	Human cDNA differe
3	2370	100.0	2370	24 ABR97360	Gene #3858 used to
4	2370	100.0	2370	24 ABL70006	Pancreas cancer re
5	2312.6	97.6	2343	25 ABR74395	Human cDNA sequenc
6	2301.6	97.1	2346	20 AAV74190	Human sgk DNA. Ho
7	2262.8	95.5	2311	19 AAV23833	Human protein kina
8	2262.8	95.5	2311	25 ACA56174	Human signalling p

ALIGNMENTS

RESULT 1
AAV48311
ID AAV48311 standard; cDNA; 2370 bp.
XX
AC AAV48311;
XX
DT 16-NOV-1998 (first entry)
XX
DE Human cell-volume regulating kinase h-sgk.
XX
KW ss; human; cell-volume; kinase; h-sgk; diabetes mellitus;
XX renal insufficiency; inflammation; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 43..1338
FT /*tag= a
FT /product= "Kinase h-sgk"
XX
PN EP861896-A2.
XX
PD 02-SEP-1998.
XX
PF 27-JAN-1998; 98EP-010138.
XX
PR 28-FEB-1997; 97DE-1008173.
XX
PA (DADE-) DADE BEHRING MARBURG GMBH.
XX
PI Lang F, Waldegger S;

Human breast tumor
Novel protein kina
DNA encoding a rat
Novel protein kina
cDNA encoding nove
Novel protein kina
Human serum and q1
Human serum and q1
Human serum and q1
Human serum and q1
Human serum glucoc
DNA encoding novel
DNA encoding novel
Human polynucleoti
Human polynucleoti
Human serine threo
Human protein enco
Human secreted pro
Human prostate ran
Nucleotide sequenc
Novel protein kina
cDNA encoding nove
DNA encoding novel
Human full-length
Human gene signatu
Bovine EST associa
Bovine EST associa
Human colon specif
Novel protein kina
Drosophila melanog
Drosophila melanog
Androgen receptor
Human Akt-3 coding
Human Akt-3 encodin
Human Akt-3 nucleo
DNA encoding a hum
Human Akt3 encodin

XX	WPI: 1998-449109/39.
DR	P-PSDB; AAW77219.
DR	
XX	
XX	New nucleic acid encoding cell-volume regulating kinase h-sgk and
PT	related proteins - used for diagnosis and treatment of diseases
PT	involving changes in cell volume, e.g. renal insufficiency,
PT	inflammation, infections etc.
XX	
XX	Disclosure: Fig 1: 15pp; German.
PS	
XX	
XX	The human cell-volume regulating kinase h-sgk is inhibited by the
CC	swelling of cells (or presence of urea), whereas cell shrinkage
CC	stimulates its expression. The nucleic acid h-sgk, and fragments, are
CC	particularly used to detect changes in cell volume, specifically for
CC	diagnosis of conditions that involve such changes, e.g. hyper- and hypo-
CC	natraemia, diabetes mellitus, renal insufficiency, hypercatabolism,
CC	hepatic encephalopathy, inflammation, microbial/viral infection, fructose
CC	intolerance, hyper- and hypo-glycaemia and Alzheimer's disease.
CC	The nucleic acid, protein and products including receptors that bind
CC	h-sgk, can be used to treat these disorders.
XX	
XX	Sequence 2370 BP: 636 A; 517 C; 513 G; 704 T; 0 other;
50	

Db	661	 GCGAGTGCCCTGGGCTACCTGCAATTCACATCAACATCGTTATATAGAGACTTAAACACAGAG	720
Qy	721	 AATAATTTGCTAGATTCACAGAGACACATTTGCTCTTACTTGATTTCSGACTCTGCAAGAG	780
Db	721	 AATAATTTGCTAGATTCACAGAGACACATTTGCTCTTACTTGATTTCSGACTCTGCAAGAG	780
Qy	781	 AACATTGAACACAACAGCACATCCACCTTCTGTGGCACGCCGAGATCATCTGCACTT	840
Db	781	 AACATTGAACACAACAGCACATCCACCTTCTGTGGCACGCCGAGATCATCTGCACTT	840
Qy	841	 GAGGTGCTTCATPAAGCAGCGCTTATGACAGAGACTGTGGACTGTGGCTGCTGAGAGTGTGTC	900
Db	841	 GAGGTGCTTCATPAAGCAGCGCTTATGACAGAGACTGTGGACTGTGGCTGCTGAGAGTGTGTC	900
Qy	901	 TTGTATGAGATGCTGTATGGCGCTGCCGCGCTTTTATATAGCCGAACACACAGCTCAAAATGTAC	960
Db	901	 TTGTATGAGATGCTGTATGGCGCTGCCGCGCTTTTATATAGCCGAACACACAGCTCAAAATGTAC	960
Qy	961	 GACAACNTCTGAACAGAGCTCTCCAGCTGTGAACCAAAATATTACAAATTCGCGACAGACAC	1020
Db	961	 GACAACATCTGAACAGAGCTCTCCAGCTGTGAACCAAAATATTACAAATTCGCGACAGACAC	1020
Qy	1021	 CTCCTGGAGGGCTCTCTGCAGAAGSAGSAGSACAAAGCGCTCGGGCCCAAGATGATCAATTC	1080
Db	1021	 CTCCTGGAGGGCTCTCTGCAGAAGSAGSAGSACAAAGCGCTCGGGCCCAAGATGATCAATTC	1080
Qy	1081	 ATCGAGATTAAAGAGTCATGTCTCTCTCTCTTAATTAAGCTGGAGATCATCTCAATTAATAAG	1140
Db	1081	 ATCGAGATTAAAGAGTCATGTCTCTCTCTCTTAATTAAGCTGGAGATCATCTCAATTAATAAG	1140
Qy	1141	 AAGATTACTCCCGCTTTTAACCCAAATGTAGTGGGCCCAACAGAGCTAGGGATATTTAGAC	1200
Db	1141	 AAGATTACTCCCGCTTTTAACCCAAATGTAGTGGGCCCAACAGAGCTAGGGATATTTAGAC	1200
Qy	1201	 CCGAGCTTTACCGAAGAGCGCTGTCCCAACTCCATTGGCGAAGTCCCTGACACAGCTCTC	1260
Db	1201	 CCGAGCTTTACCGAAGAGCGCTGTCCCAACTCCATTGGCGAAGTCCCTGACACAGCTCTC	1260
Qy	1261	 GTCACAGCCAGCGCTCAAGSAGCTGCCGAGCGCTTCTTAGGCTTTTCTATGCGCTGCC	1320
Db	1261	 GTCACAGCCAGCGCTCAAGSAGCTGCCGAGCGCTTCTTAGGCTTTTCTATGCGCTGCC	1320
Qy	1321	 ACGGACTTTTCTCTCTGAACCCCTGTTAGGCGCTTGGTTTAAAGATTTTATGTGTGTCTTC	1380
Db	1321	 ACGGACTTTTCTCTCTGAACCCCTGTTAGGCGCTTGGTTTAAAGATTTTATGTGTGTCTTC	1380
Qy	1381	 CGAATGTTTAGTTAGCCCTTTTGTGTAGCGCGCGAGCTGACAGAGATCTTTACAGAGAA	1440
Db	1381	 CGAATGTTTAGTTAGCCCTTTTGTGTAGCGCGCGAGCTGACAGAGATCTTTACAGAGAA	1440
Qy	1441	 TTTGACATCTCTGGAAGCTTAGTAATCTTAFTGACACAGCTTCCGCTGGAAATTTTGTAA	1500
Db	1441	 TTTGACATCTCTGGAAGCTTAGTAATCTTAFTGACACAGCTTCCGCTGGAAATTTTGTAA	1500
Qy	1501	 GAGCACATCTCCTAGTGAGCTCATAGAGTTTTCATTTTATCTTCTCTTCCAAAGCTGG	1560
Db	1501	 GAGCACATCTCCTAGTGAGCTCATAGAGTTTTCATTTTATCTTCTCTTCCAAAGCTGG	1560
Qy	1561	 TGCTATCTCTGAACAGCGCTTAGGTGCGCGCTTAGACGAGCGCAGGAGCTTTCGCTTAGA	1620
Db	1561	 TGCTATCTCTGAACAGCGCTTAGGTGCGCGCTTAGACGAGCGCAGGAGCTTTCGCTTAGA	1620
Qy	1621	 AAGCGGACCTGTTCTTAAAAAGGCTCTCTGCAGATCTGTCTGGGCTGTGATGACGAAAT	1680
Db	1621	 AAGCGGACCTGTTCTTAAAAAGGCTCTCTGCAGATCTGTCTGGGCTGTGATGACGAAAT	1680
Qy	1681	 TATCAAAATGTGCCCTTTCTGAAGAGATGTGTTAGCTCCAAAGCTTTTCTATCGCAGTG	1740
Db	1681	 TATCAAAATGTGCCCTTTCTGAAGAGATGTGTTAGCTCCAAAGCTTTTCTATCGCAGTG	1740
Qy	1741	 TTTCAGTCTTTATTTTCCCTTTGTGGAATATGCTCTGTGAACCGCTCGTGTGAGTGTGCTAT	1800

Db 241 GAGCTTATGAATGCCAACCTTCTCTCCACCAAGTCTCTCTCAGCAAAATCAACCTTGGC 300
QY 301 CCGTCGTCATCCATCTGCTAAACCATCTCAGCTTCTCACTCTCTCAAGTATCGGAAG 360
Db 301 CCGTCGTCATCCATCTGCTAAACCATCTCAGCTTCTCACTCTCTCAAGTATCGGAAG 360
QY 361 GCGAGTTTGGAAAGGTTCTTCTAGCAGACACAGGAGGAGAGTGTCTATGCGATC 420
Db 361 GCGAGTTTGGAAAGGTTCTTCTAGCAGACACAGGAGGAGAGTGTCTATGCGATC 420
QY 421 AAAGTTTACAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGATATATGTCGAG 480
Db 421 AAAGTTTACAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGATATATGTCGAG 480
QY 481 CGGAATGTTCTGTTGAAGAATGTTGAAGCACCCTTCTCTGTTGGGCTTCACTTCTTTC 540
Db 481 CGGAATGTTCTGTTGAAGAATGTTGAAGCACCCTTCTCTGTTGGGCTTCACTTCTTTC 540
QY 541 CAGACTGCTGACAAATTTGCTGCTAGACTACATTAATGTTGGAGAGTGTGTTCTAC 600
Db 541 CAGACTGCTGACAAATTTGCTGCTAGACTACATTAATGTTGGAGAGTGTGTTCTAC 600
QY 601 CATCTCCAGAGGAAACGCTGCTCTGGAACACCGGCTGCTTCTATGCTGCTGAAATA 660
Db 601 CATCTCCAGAGGAAACGCTGCTCTGGAACACCGGCTGCTTCTATGCTGCTGAAATA 660
QY 661 GCCAGTGGCTGGGCTACTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 720
Db 661 GCCAGTGGCTGGGCTACTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 720
QY 721 AATATTTGCTAGATTCACAGGACACATCTGCTGACGCGGCTGCTGACGCTGATCTG 780
Db 721 AATATTTGCTAGATTCACAGGACACATCTGCTGACGCGGCTGCTGACGCTGATCTG 780
QY 781 AACATTTGAACACACAGCACAACATCTGCTGACGCGGCTGCTGACGCTGATCTG 840
Db 781 AACATTTGAACACACAGCACAACATCTGCTGACGCGGCTGCTGACGCTGATCTG 840
QY 841 GAGTGTCTCATAGCAGCCTTATGACAGGACTGTGAGTGTGCTGGAGCTGTC 900
Db 841 GAGTGTCTCATAGCAGCCTTATGACAGGACTGTGAGTGTGCTGGAGCTGTC 900
QY 901 TTGATGATGCTGTATGGCTGCGGCTTTTATAGCCGAACACACAGCTGAATGTAC 960
Db 901 TTGATGATGCTGTATGGCTGCGGCTTTTATAGCCGAACACACAGCTGAATGTAC 960
QY 961 GACACATTTGACACAGCCTCTCAGAGGACAGGACAGGCTGACAAATTCGCGAAGAC 1020
Db 961 GACACATTTGACACAGCCTCTCAGAGGACAGGACAGGCTGACAAATTCGCGAAGAC 1020
QY 1021 CTCTGGAGGCTCTCTCAGAGGACAGGACAGGCTGCGGCTGCGGCTGCGGCTGCTC 1080
Db 1021 CTCTGGAGGCTCTCTCAGAGGACAGGACAGGCTGCGGCTGCGGCTGCGGCTGCTC 1080
QY 1081 ATGGAGATTAAGAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Db 1081 ATGGAGATTAAGAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1141 AAGATTACTCCCTTTTAAACCAATGTGAGTGGGCTGAGGCTGAGGCTGAGGCTTTCAC 1200
Db 1141 AAGATTACTCCCTTTTAAACCAATGTGAGTGGGCTGAGGCTGAGGCTGAGGCTTTCAC 1200
QY 1201 CCGGAGTTTACGAGAGCCTGCTCCCAACTCTGATGGCAAGTCCCTGACAGGCTGCTC 1260
Db 1201 CCGGAGTTTACGAGAGCCTGCTCCCAACTCTGATGGCAAGTCCCTGACAGGCTGCTC 1260
QY 1261 GTCACAGCCAGGCTGAGGAGCTGCGGAGGCTTCTCTAGGCTTCTCTAGGCTTCTCTAG 1320
Db 1261 GTCACAGCCAGGCTGAGGAGCTGCGGAGGCTTCTCTAGGCTTCTCTAGGCTTCTCTAG 1320
QY 1321 ACGGACTCTTCTCTGACACCTGTTAGGCTTGTGTTTAAAGGATTTATGTTGTTTC 1380
Db 1321 ACGGACTCTTCTCTGACACCTGTTAGGCTTGTGTTTAAAGGATTTATGTTGTTTC 1380

RESULT 3
ABN97360
ID ABN97360 standard; DNA; 2370 BP.

QY 1381 CGAATGTTTAGTTAGCTTTTGGTGGAGCCGAGCTGACGAGACATCTTACAGAGAA 1440
Db 1381 CGAATGTTTAGTTAGCTTTTGGTGGAGCCGAGCTGACGAGACATCTTACAGAGAA 1440
QY 1441 TTTCACATCTCTCGAAGCTTAGCAATCTTATTCACACTGTTCTGCTGGAATTTTTCAA 1500
Db 1441 TTTCACATCTCTCGAAGCTTAGCAATCTTATTCACACTGTTCTGCTGGAATTTTTCAA 1500
QY 1501 GAGCACATCTCTCAGTGAATCTATGAGGTTTTCATTTTATCTTCTCTTCCCAACGTTG 1560
Db 1501 GAGCACATCTCTCAGTGAATCTATGAGGTTTTCATTTTATCTTCTCTTCCCAACGTTG 1560
QY 1561 TGCTATCTCTGAAACGAGCGTTAGAGTCCGCCCTTAGACGAGGAGGAGTTCGTTAGA 1620
Db 1561 TGCTATCTCTGAAACGAGCGTTAGAGTCCGCCCTTAGACGAGGAGGAGTTCGTTAGA 1620
QY 1621 AAGCGGACCTCTCTTAAAAAGGCTCTCTGAGATCTGCTGGGCTGTGATGACGAATAT 1680
Db 1621 AAGCGGACCTCTCTTAAAAAGGCTCTCTGAGATCTGCTGGGCTGTGATGACGAATAT 1680
QY 1681 TATGAATGTCGCTTTTCTGAAGAGATTTGCTAGCTCCAAAGCTTTTCTTATGCACTG 1740
Db 1681 TATGAATGTCGCTTTTCTGAAGAGATTTGCTAGCTCCAAAGCTTTTCTTATGCACTG 1740
QY 1741 TTTCAATGTTCTTTTATTTTCCCTTGTGGATATGCTGTGAACTGCTGTGAGTGTGAT 1800
Db 1741 TTTCAATGTTCTTTTATTTTCCCTTGTGGATATGCTGTGAACTGCTGTGAGTGTGAT 1800
QY 1801 GCCTGATCAGAGATGATTTTGTATAGCAATCAATGTGACACTTGCAGGACATTAAC 1860
Db 1801 GCCTGATCAGAGATGATTTTGTATAGCAATCAATGTGACACTTGCAGGACATTAAC 1860
QY 1861 GTGGACATCTGTTCTTCTTCCATATTTTGAAGATAAATTTATGTTAGTGTGATGTTTGT 1920
Db 1861 GTGGACATCTGTTTCTTCCATATTTTGAAGATAAATTTATGTTAGTGTGATGTTTGT 1920
QY 1921 AAGTACGCTTAACTAACTTAAATTTTGAATGCTTGTGAATGCTGCTGATGATGATG 1980
Db 1921 AAGTACGCTTAACTAACTTAAATTTTGAATGCTTGTGAATGCTGCTGATGATGATG 1980
QY 1981 CCTAAGAAAGCAATGCTGCTACAAATTTTCTATTTTGAAGAGGTTTATGAGGACCA 2040
Db 1981 CCTAAGAAAGCAATGCTGCTACAAATTTTCTATTTTGAAGAGGTTTATGAGGACCA 2040
QY 2041 ATGCCCCAGTTGTCAGTCAAGCGCTGCTGCTTTTCTATTTTAAATGTCAGCTGTAA 2100
Db 2041 ATGCCCCAGTTGTCAGTCAAGCGCTGCTGCTTTTCTATTTTAAATGTCAGCTGTAA 2100
QY 2101 AATGGGCAATTTATGTTTATTTTTCATCTCTGATATTTGATGATGATGATGATG 2160
Db 2101 AATGGGCAATTTATGTTTATTTTTCATCTCTGATATTTGATGATGATGATGATG 2160
QY 2161 AACGCTGTACATTTGGTTTAAACACTAGTATATTTTAACTTACAGGCTTATTTGATG 2220
Db 2161 AACGCTGTACATTTGGTTTAAACACTAGTATATTTTAACTTACAGGCTTATTTGATG 2220
QY 2221 TAAACCACTTAACTGCTACTGTAATTAACATGTTTATATAGTCAATGCTTCTCTC 2280
Db 2221 TAAACCACTTAACTGCTACTGTAATTAACATGTTTATATAGTCAATGCTTCTCTC 2280
QY 2281 ATCCCATCAGACACTTTTGTGCTGATTAACACTGATTTTGGTTTGCATTAACACT 2340
Db 2281 ATCCCATCAGACACTTTTGTGCTGATTAACACTGATTTTGGTTTGCATTAACACT 2340
QY 2341 GAAAAATAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2370
Db 2341 GAAAAATAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2370

XX ABN97360;
AC
XX
XX
XX
XX
DE Gene #3858 used to diagnose liver cancer.
XX
XX Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
OS
XX
XX W0200229103-A2.
XX
XX 1J-APK-2002.
XX
XX 02-OCT-2001; 2001WO-US30589.
XX
XX 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENH LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
PI WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient, in a
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample.
XX
XX Claim 1; SEQ ID NO 3858; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 2370 BP; 636 A; 517 C; 513 G; 704 T; 0 other;
XX
XX Query Match 100.0%; Score 2370; DB 24; Length 2370;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CACGAGGAGCGCTAACGCTTTCTCTCCCGCGGTGTCATGACGGTGAACACTGAG 60
XX
XX 1 CACGAGGAGCGCTAACGCTTTCTCTCCCGCGGTGTCATGACGGTGAACACTGAG 60
XX
XX 61 GCTGCTAAGGGACCCCTACCTTACTCCAGGATGAGGGGCATGGTGGCAATCTCATCGCT 120
XX
XX 61 GCTGCTAAGGGACCCCTACCTTACTCCAGGATGAGGGGCATGGTGGCAATCTCATCGCT 120
XX
XX 121 TTCTATGAAGCAGGAGGAGTGGTCTGAACGACTTTATTTCAGAGATGGCAATAACTCC 180
XX
XX 121 TTCTATGAAGCAGGAGGAGTGGTCTGAACGACTTTATTTCAGAGATGGCAATAACTCC 180
XX
XX 181 TATGATGCAACACCCCTGAGTTCAGTCCATCTTTGAAGATCTCCCAACCTCAGAGCGCT 240
XX
XX 181 TATGATGCAACACCCCTGAGTTCAGTCCATCTTTGAAGATCTCCCAACCTCAGAGCGCT 240
XX
XX 241 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCTCAGCAATCAACCTTGGC 300

DB 241 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCTCAGCAATCAACCTTGGC 300
QY 301 CGTGTGTCCTCAATCTTCATGCAAAACCACTTTCACCTTCTTGAAGTGAACGAAAG 160
DB 301 CGTGTGTCCTCAATCTTCATGCAAAACCACTTTCACCTTCTTGAAGTGAACGAAAG 160
QY 361 GGCAGTCTTGAAGAGTCTTCTTAGCAAGACACAAAGGAGAGAGTGTCTTAAGAGTC 420
DB 361 GGCAGTCTTGAAGAGTCTTCTTAGCAAGACACAAAGGAGAGAGTGTCTTAAGAGTC 420
QY 421 AAGTCTTACAGAGAGAGAGTCTTCTTAGCAAGAGAGAGAGAGAGTGTCTTAAGAGTC 480
DB 421 AAGTCTTACAGAGAGAGAGTCTTCTTAGCAAGAGAGAGAGAGAGTGTCTTAAGAGTC 480
QY 481 CGGAATGTTCTGTGTAAGAAATCTGAAGACACCTTTCCTGTGGGCGCTTCTCTTTC 540
DB 481 CGGAATGTTCTGTGTAAGAAATCTGAAGACACCTTTCCTGTGGGCGCTTCTCTTTC 540
QY 541 CAGACTGCTGACAAATCTGACTTCTCTAGCTACATTAATGCTGGAGAGTGTCTTCTAC 600
DB 541 CAGACTGCTGACAAATCTGACTTCTCTAGCTACATTAATGCTGGAGAGTGTCTTCTAC 600
QY 601 CATCTCCAGAGGAGAGTCTTCTCTGAAACACAGGAGTCTTCTTCTATGCTGTCTTCT 660
DB 601 CATCTCCAGAGGAGAGTCTTCTCTGAAACACAGGAGTCTTCTTCTATGCTGTCTTCT 660
QY 661 GGCAGTGTCTTGGGCTACCTGATCTGACTGACATCTGCTTATAGAGACTTAAAGTAC 720
DB 661 GGCAGTGTCTTGGGCTACCTGATCTGACTGACATCTGCTTATAGAGACTTAAAGTAC 720
QY 721 AATATTTGCTAGATTTCACAGGAGAGTCTTCTCTTCTGACTTCTGAGAGTCTTCT 780
DB 721 AATATTTGCTAGATTTCACAGGAGAGTCTTCTCTTCTGACTTCTGAGAGTCTTCT 780
QY 781 ACATTTGAACACACAGAGACACATCTTCTGTCGACGCGCGGAGTATCTTCTGACCT 840
DB 781 ACATTTGAACACACAGAGACACATCTTCTGTCGACGCGCGGAGTATCTTCTGACCT 840
QY 841 GAGGTGCTTCTAAGCAGCGCTTATGACAGGAGTCTTCTGAGTCTGCTGCTGAGTCT 900
DB 841 GAGGTGCTTCTAAGCAGCGCTTATGACAGGAGTCTTCTGAGTCTGCTGCTGAGTCT 900
QY 901 TTGTATGAGATGCTGATGAGTCTGCTGCTGCTTCTTATAGCGGAGACAGTCTTCT 960
DB 901 TTGTATGAGATGCTGATGAGTCTGCTGCTGCTTCTTATAGCGGAGACAGTCTTCT 960
QY 961 GACAACATTTCTGAACAGAGTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1020
DB 961 GACAACATTTCTGAACAGAGTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1020
QY 1021 CTCTCGAGGAGGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 CTCTCGAGGAGGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 ATGGAGATTAAAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
DB 1081 ATGGAGATTAAAGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
QY 1141 AAGATTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
DB 1141 AAGATTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200
QY 1201 CCGAGTCTTACCGAG 1260
DB 1201 CCGAGTCTTACCGAG 1260
QY 1261 GTCACAG 1320
DB 1261 GTCACAG 1320
QY 1321 AGGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380
DB 1321 AGGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380

Db 1321 ACGGACTCTTCTCTGACCCCTGTTAGGCTTGGTTTTAAAGGATTTTATGTCTGTTTC 1380
 Qy 1381 CGATGTTTTAGTAGCTTTGTTGGAGCCGCCAGCTGACGACATCTTACAGAGAA 1440
 Db 1381 CGAATGTTTTAGTAGCTTTGTTGGAGCCGCCAGCTGACGACATCTTACAGAGAA 1440
 Qy 1441 TTTCACATCTCTGGAAGCTTACCAATCTTATGACACACTGTCGCTGGAATTTTGA 1500
 Db 1441 TTTCACATCTCTGGAAGCTTACCAATCTTATGACACACTGTCGCTGGAATTTTGA 1500
 Qy 1501 GAGCAGATCTCTCAGTCAGCTCATGAGTTTTCATTTTATCTTCTTCCCAACGTGG 1560
 Db 1501 GAGCAGATCTCTCAGTCAGCTCATGAGTTTTCATTTTATCTTCTTCCCAACGTGG 1560
 Qy 1561 TGTATCTCTGAAACGAGCGTTAGAGTCCGCTTTAGACGAGGAGGAGTTTCGTTAGA 1620
 Db 1561 TGTATCTCTGAAACGAGCGTTAGAGTCCGCTTTAGACGAGGAGGAGTTTCGTTAGA 1620
 Qy 1621 AAGCGACCTGTTCTAAAAAGCTCTCCCTGACAGATCTGCTGGGCTGTGATGACGAATAT 1680
 Db 1621 AAGCGACCTGTTCTAAAAAGCTCTCCCTGACAGATCTGCTGGGCTGTGATGACGAATAT 1680
 Qy 1681 TATGAATGTCCTTTCTGAGAGATCTGTTAGCTCCAAAGCTTTTCTTATCGACGTG 1740
 Db 1681 TATGAATGTCCTTTCTGAGAGATCTGTTAGCTCCAAAGCTTTTCTTATCGACGTG 1740
 Qy 1741 TTTCAGTCTTTATTTTCCCTTGTGATATGCTGTGTAACCTGCTGTGAGTGTGTTAT 1800
 Db 1741 TTTCAGTCTTTATTTTCCCTTGTGATATGCTGTGTAACCTGCTGTGAGTGTGTTAT 1800
 Qy 1801 GCCTGATCAGATGATTTGTTTATGACATCAATGTGACACTTTCAGGACACATCAAC 1860
 Db 1801 GCCTGATCAGATGATTTGTTTATGACATCAATGTGACACTTTCAGGACACATCAAC 1860
 Qy 1861 GTGGGACATTTGTTTCTTCCATATTTGGAGATATAATTTATGTTGAGACTTTTGT 1920
 Db 1861 GTGGGACATTTGTTTCTTCCATATTTGGAGATATAATTTATGTTGAGACTTTTGT 1920
 Qy 1921 AAGTACGGTTAATACTAAATTTATGAAATGGCTTGGCAATGACTCGTATTCAGATG 1980
 Db 1921 AAGTACGGTTAATACTAAATTTATGAAATGGCTTGGCAATGACTCGTATTCAGATG 1980
 Qy 1981 CCTAAGAAAGCATGCTGCTACAAATATTTCTATTTTATGAAAGGGTTTTATGACCA 2040
 Db 1981 CCTAAGAAAGCATGCTGCTACAAATATTTCTATTTTATGAAAGGGTTTTATGACCA 2040
 Qy 2041 ATGCCCGAGTTGTCAGTCAGCCGCTGGTGTGTTTTCATTTTAAATGTCACCTGTAA 2100
 Db 2041 ATGCCCGAGTTGTCAGTCAGCCGCTGGTGTGTTTTCATTTTAAATGTCACCTGTAA 2100
 Qy 2101 AATGGCATTTATTTATGTTTTTTTTCATTTCTGATATTTATGTTATGTTATAAG 2160
 Db 2101 AATGGCATTTATTTATGTTTTTTTTCATTTCTGATATTTATGTTATGTTATAAG 2160
 Qy 2161 AAGCTGTGATGTTGTTTATACACTAGTATATTTAACTTACAGGCTTATTTGTAATG 2220
 Db 2161 AAGCTGTGATGTTGTTTATACACTAGTATATTTAACTTACAGGCTTATTTGTAATG 2220
 Qy 2221 TAAACCCACCTTTTATGTTACTGTATTAACATGGTTATATACGTCACATCTTCCCTC 2280
 Db 2221 TAAACCCACCTTTTATGTTACTGTATTAACATGGTTATATACGTCACATCTTCCCTC 2280
 Qy 2281 ATCCCTCAGACACACTTTTGTGTGTTGATTAACATGATTTTGGTTTGCATTAACCTTT 2340
 Db 2281 ATCCCTCAGACACACTTTTGTGTGTTGATTAACATGATTTTGGTTTGCATTAACCTTT 2340
 Qy 2341 GAAAT 2370
 Db 2341 GAAAT 2370

RESULT 4
 ABL70006

ID XX
 AC XX
 XX XX
 DT 15-MAY-2002 (first entry)
 XX DE
 DE DE
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX Homo sapiens.
 OS WO200194629-A2.
 XX 13-DEC-2001.
 XX 30-MAY-2001; 2001WO-US10838.
 XX 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 23-SEP-2000; 2000US-234923P.
 PR 23-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 28-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horriqan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and

XX
PS

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 2370 BP: 636 A; 517 C; 513 G; 704 T; 0 other;
XX
50

Query Match	100.0%	Score 2370;	DB 24;	Length 2370;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2370;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	CACGAGGAGGCGTAAACGTTCTTCTGTCTCCCGCGGTGTGTATGACGCTGAAAACACTGAG	60		
DB	1	CACGAGGAGGCGTAAACGTTCTTCTGTCTCCCGCGGTGTGTATGACGCTGAAAACACTGAG	60		
QY	61	GCTGCTAAGGCGACCCCTCACTTACTTCAGGATGAGGGGCATGGTGGCAATTCCTCATCGCT	120		
DB	61	GCTGCTAAGGCGACCCCTCACTTACTTCAGGATGAGGGGCATGGTGGCAATTCCTCATCGCT	120		
QY	121	TTCTAAGAAGCAGAGGAGGATGGGTGTGAAGCACTTTATTTCAGAAGATTGCCAATAAATCC	180		
DB	121	TTCTATGTAAGCAGAGGAGGATGGGTGTGAAGCACTTTATTTCAGAAGATTGCCAATAAATCC	180		
QY	181	TATGCATGCAACACCCCTGAAGTTCAGTCCCATCTTGAAGATCTCCCAACCTTCAGGAGCT	240		
DB	181	TATGCATGCAACACCCCTGAAGTTCAGTCCCATCTTGAAGATCTCCCAACCTTCAGGAGCT	240		
QY	241	GAGCTTATGAATGCCAACCCCTCTCTCCACCAAGTCTTCTTCAGCAAAATCAACCTTGGC	300		
DB	241	GAGCTTATGAATGCCAACCCCTCTCTCCACCAAGTCTTCTTCAGCAAAATCAACCTTGGC	300		
QY	301	CGCTCGTCCCATCTCATGTAAACCATCTGACTTTCATCTTCTTGAAGTGTATCGGAAG	360		
DB	301	CGCTCGTCCCATCTCATGTAAACCATCTGACTTTCATCTTCTTGAAGTGTATCGGAAG	360		
QY	361	GGCAGTTTGTGAAAGGTTCTTCTAGCAAGACACAGGAGAGAGTGTTCATCGACGTC	420		
DB	361	GGCAGTTTGTGAAAGGTTCTTCTAGCAAGACACAGGAGAGAGTGTTCATCGACGTC	420		
QY	421	AAAGTTTTCAGAGAGAAGCAATCCTGAAAAGAAAGAGAGAGCATATTATGTGCGAG	480		
DB	421	AAAGTTTTCAGAGAGAAGCAATCCTGAAAAGAAAGAGAGAGCATATTATGTGCGAG	480		
QY	481	CGGAATGTTCTGTGGAAGATGTGAGACACCCCTTCTCTGTGGGCCCTTCACATTCCTTTC	540		
DB	481	CGGAATGTTCTGTGGAAGATGTGAGACACCCCTTCTCTGTGGGCCCTTCACATTCCTTTC	540		
QY	541	CAGACTGCTGACAAATTTGATCTTGTCTCTAGACTACATTAATGGTGGAGAGTTGTCTAC	600		
DB	541	CAGACTGCTGACAAATTTGATCTTGTCTCTAGACTACATTAATGGTGGAGAGTTGTCTAC	600		
QY	601	CATCTCCAGAGGAGCGCTGCTTCTTGGACCAAGGCGTCTTCTATCTGCTCTGAAATA	660		
DB	601	CATCTCCAGAGGAGCGCTGCTTCTTGGACCAAGGCGTCTTCTATCTGCTCTGAAATA	660		
QY	661	GCCAGTGCCCTGGGCTACCTGCATCTCACTGAACATCGTTTATAGAGACTTAAACACGAG	720		

1741	Db	TTTCAGTCTCTTATTTTCCCTTGTGGGATATGCTGTGTGAACCGTCGTCGTGAGTGTGGTAT	1800
1801	QY	GCCTGATACAGATGGATTTTGTATTAAGCATCAATGTGACACVTGAGGACACTACAAC	1860
1801	Db	GCCTGATACAGATGGATTTTGTATTAAGCATCAATGTGACACTTGCAGGACACTACAAC	1860
1861	QY	GTGGGACATTTGTGTCTCCATATTTGGAAGATAAATTTATGCTAGACATTTTGTG	1920
1861	Db	GTGGGACATTTGTGTCTCCATATTTGGAAGATAAATTTATGCTAGACATTTTGTG	1920
1921	QY	AAGATCGGTTAATAAATACTAAAATTTATGAAATGCTTTCGAAATGACTCGTATTCAGATG	1980
1921	Db	AAGATCGGTTAATAAATACTAAAATTTATGAAATGCTTTCGAAATGACTCGTATTCAGATG	1980
1981	QY	CCTAAGCAAAACAGTTCGCTGCTACAATATTTCTATTTTGAAGAAGGTTTTTATGGACCA	2040
1981	Db	CCTAAGCAAAAGCATTCGCTGCTACAATATTTCTATTTTGAAGAAGGTTTTTATGGACCA	2040
2041	QY	ATGCCACAGTTGTCAGTCAGACGCGTTTGCTGTTTTTCATGTTTAAATATGTCACCTGTAA	2100
2041	Db	ATGCCACAGTTGTCAGTCAGACGCGTTTGCTGTTTTTCATGTTTAAATATGTCACCTGTAA	2100
2101	QY	AATGGCATTATTTATGTTTTTTTTTTCGATTCCTGATAATTTGTATGTATGTATAAG	2160
2101	Db	AATGGCATTATTTATGTTTTTTTTTTCGATTCCTGATTAATTTGTATGTATGTATAAG	2160
2161	QY	AACGTCGTACATTTGGTTATACACTAGTATATTTAAACTTCAGGCTTATTTGTAATG	2220
2161	Db	AACGTCGTACATTTGGTTATACACTAGTATATTTAAACTTCAGGCTTATTTGTAATG	2220
2221	QY	TAAACCACCATTTTAAATGTAAGTAAATTAACATGGTTATTAATAGCTACAATCTTCCTC	2280
2221	Db	TAAACCACCATTTTAAATGTAAGTAAATTAACATGGTTATTAATAGCTACAATCTTCCTC	2280
2281	QY	ATCCCATACACAACTTTTTTGTGTGTGATAAACTGATTTTGGTTGCCAATAAACCTT	2340
2281	Db	ATCCCATACACAACTTTTTTGTGTGTGATAAACTGATTTTGGTTGCCAATAAACCTT	2340
2341	QY	GAAGAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	2370
2341	Db	GAAGAAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA	2370

845117 5

RESOLVED
50V LAY 305

ABX/4395
10 ABX74395 standard: CDNA: 2343 BP.

yy
ID
ABX/4395

XX
XC
APV7A395:

AC
VY
ABX/4395;

XX
DT 21-MAR-2003 (first entry)

DT 21-MAR-2003 (first entry)

[illegible]

DE: Human cDNA sequence #0 up-regulated in non-aggressive

XX
XX

KW Human; microarray
816097

KW differential
 KW non-differential
 KW

non-aggressive CC-RCC tumours

gene expression profiling; t

[illegible]

Homo sapiens.

XX
DN
W030030411-23

XX

New microarray, comprising a matrix of cDNA probe from a set of probes immobilised to a solid surface in predetermined order, useful in the prognosis of patients with clear cell renal carcinoma.

PROSECUTOR OR DEFENDANT

Claim 1; Page 79-80; 179pp; English.

The present invention relates to a microarray comprising a matrix of at least one cDNA probe from a set of probes immobilised to a solid surface in a predetermined order, where a row of pixels corresponds to replicates of one distinct probe from the set. The probes are complementary to nucleic acid sequences that are expressed differentially in aggressive as compared to non-aggressive types of clear cell renal carcinoma (CC-RCC) and that hybridise to the probes under high stringency conditions. The microarray is useful for the prognosis of patients with CC-RCC, wherein aggressive and non-aggressive CC-RCC tumour types are characterised by differential expression profiles of genes that hybridise with one or more probes immobilised on the microarray. The arrays are useful for gene expression profiling of tumour and normal tissues. The present sequence represents a human cDNA sequence up-regulated in non-aggressive CC-RCC phenotypes.

Recommended by _____

Sequence 2343 BP; 612 A; 517 C; 511 G; 703 T; 0 other;
 ary Match 97 64; Score 2312 6; DB 25; Length 2143;

Query Match	99.78%	Score 2314266	DB 237	Length 23479
Best Local Similarity	99.78%	Pred. No. 0:		
Matches 2337;	Conservative 0:	Mismatches 4:	Indels 2:	Gaps 2
QY	8	GAGCGCTAACGCTTTTCTGTCTCCCGCGGTGGTGATGACGGTGAAGACTGAGGCTGCTA	67	
DB	1	GAGCGCTAACGCTTTTCTGTCTCCCGCGGTGGTGATGACGGTGAAGACTGAGGCTGCTA	60	
QY	68	AGGGCACCCCTCACTTACTCCAGGATGAGGGGCATGTGGCAATTCATCGCTTTCATGA	127	
DB	61	AGGGCACCCCTCACTTACTCCAGGATGAGGGGCATGTGGCAATTCATCGCTTTCATGA	120	
QY	128	ACGAGAGCAGATGGCTCTGCAACCACTTTATTCAGAGAGATTGCCAATAAGCTCCTATGCAAT	187	
DB	121	ACGAGAGGAGATGGGCTCGAACCACTTTATTCAGAGAGATTGCCAATAAGCTCCTATGCAAT	180	
QY	188	GCACACACCCCTGAAGTTCAGTCCATCTTTGAAGATCTCCCAACCTCAGGAGCTTGAGCTTGA	247	
DB	181	GCACACACCCCTGAAGTTCAGTCCATCTTTGAAGATCTCCCAACCTCAGGAGCTTGAGCTTGA	240	
QY	248	TGAATGCCAACCCCTCTCCYCCACCAAGTCCTTCTCAGCAAAATCAACCTTGGCCCGCTCGCT	307	
DB	241	TGAATGCCAACCCCTCTCCTCCACCAAGTCCTTCTCAGCAAAATCAACCTTGGCCCGCTCGCT	300	
QY	308	CCAATCCTCTACTCTAAACCATCTGACTTTTCATTTCTTGAAGATGATCGGAAGGGGAGTT	367	
DB	301	CCAATCCTCTACTCTAAACCATCTGACTTTTCATTTCTTGAAGATGATCGGAAGGGGAGTT	360	

Serum glucocorticoid regulated kinase; sgk; human; treatment: inhibitor;
 KW
 serine/threonine protein kinase family; antagonist; diabetic nephropathy;
 KW
 chronic renal failure; inflammation; Alzheimer's disease; wound; ss.
 KW

XX (SMK) SMITHKLINE BEECHAM CORP.
 XX PA Kumar JM;
 XX PI WPI: 1999-047627/05.
 XX DR P-PSDB; AAW90139.
 XX PT Treating chronic renal failure, diabetic nephropathy and Alzheimer's
 XX PT disease - by administration of nucleic acids and antagonists which
 XX PT inhibit activity or expression of human serum glucocorticoid
 XX PT regulated kinase (sgk), a serine/threonine protein kinase
 XX PS Disclosure: Page 14-15; 17pp; English.
 XX CC This sequence encodes a novel human serum glucocorticoid regulated kinase
 XX CC (sgk) protein which is a member of the serine/threonine protein kinase
 XX CC family. This protein is used for the treatment of a subject having need
 XX CC to inhibit/antagonise activity or expression of human sgk polypeptide
 XX CC e.g. for the treatment of chronic renal failure, diabetic nephropathy,
 XX CC inflammation, Alzheimer's disease and wounds.
 XX SQ Sequence 2346 BP; 608 A; 520 C; 518 G; 700 T; 0 other;
 Query Match 97.1%; Score 2301.6; DB 20; Length 2346;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2337; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
 QY 1 CACGAGGAGCGCTAACGCTTCTTCTGCTCCCGCGGTGGTGATGACGGTGAACACTGAG 60
 DB 1 CACGAGGAGCGCTAACGCTTCTTCTGCTCCCGCGGTGGTGATGACGGTGAACACTGAG 62
 QY 61 GCTGCTAAGCGACCCCTCACTTACTCCAGGATGAGGCGCATGGTGGCAATTCATCGCT 120
 DB 63 GCTGCTAAGCGACCCCTCACTTACTCCAGGATGAGGCGCATGGTGGCAATTCATCGCT 122
 QY 121 TTCATGAACAGAGGAGGAGTGGTGTGAACGACTTTATTCAGAGATGTCGAATTAATTC 180
 DB 123 TTCATGAACAGAGGAGGAGTGGTGTGAACGACTTTATTCAGAGATGTCGAATTAATTC 182
 QY 181 TATGATGCAACACACCCCTCAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAGCGT 240
 DB 183 TATGATGCAACACACCCCTCAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAGCGT 242
 QY 241 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTTCTCAGCAATCAACCTTGGC 300
 DB 243 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTTCTCAGCAATCAACCTTGGC 302
 QY 301 CGGTCTCCCAATCCTCATGCTAAACCATCTGACTTTCATCTTGAAGATGATCGGAAG 360
 DB 303 CGGTCTCCCAATCCTCATGCTAAACCATCTGACTTTCATCTTGAAGATGATCGGAAG 362
 QY 361 GCGAGTTCGGAAGTCTTCTAGCAGACACAGGCGCAGAGAGTGTCTATGCGATC 420
 DB 363 GCGAGTTCGGAAGTCTTCTAGCAGACACAGGCGCAGAGAGTGTCTATGCGATC 422
 QY 421 AAGTTCATACAGAAAGCAATCTGAAAAGAGAGAGAGCAATATATGTCGGAG 480
 DB 423 AAGTTCATACAGAAAGCAATCTGAAAAGAGAGAGAGCAATATATGTCGGAG 482
 QY 481 CGGAATGTCCTTGTGAAGATGTGAGACACCCCTTCTGCGTGGGCGCTTCACTTCCTTC 540
 DB 483 CGGAATGTCCTTGTGAAGATGTGAGACACCCCTTCTGCGTGGGCGCTTCACTTCCTTC 542
 QY 541 CAGACTCTGACAAATGTACTTGTCTTACACTACATTAATGTTGGAGAGTGTCTTAC 600
 DB 543 CAGACTCTGACAAATGTACTTGTCTTACACTACATTAATGTTGGAGAGTGTCTTAC 602
 QY 601 CATCTCCAGAGGAGCGCTTCTCTGGAACACAGCGGCTCGTCTTATGCTGCTGAAATA 660
 DB 603 CATCTCCAGAGGAGCGCTTCTCTGGAACACAGCGGCTCGTCTTATGCTGCTGAAATA 662
 QY 661 GCCAGTCTGGGCTACCTGCAATTCACCTGAACATCGTTTATAGAGACTTAAACACGAG 720
 DB 663 GCCAGTCTGGGCTACCTGCAATTCACCTGAACATCGTTTATAGAGACTTAAACACGAG 722
 QY 721 AATATTTCTGCTAGATTCACAGGACACACATTCCTTACTGATTCGGACTCTGTAAGAG 780
 DB 723 AATATTTCTGCTAGATTCACAGGACACACATTCCTTACTGATTCGGACTCTGTAAGAG 782
 QY 781 AACATTGAACACAGACACACATTCCTTCTGTCGACGCCCGAGTATATATAT 840
 DB 783 AACATTGAACACAGACACACATTCCTTCTGTCGACGCCCGAGTATATATAT 842
 QY 841 GAGGTCTCTAT 900
 DB 843 GAGGTCTCTAT 902
 QY 901 TTGATGAGATGCTGTATGCGCTGCGCGCTTTTATAGCCGAAACACAGCTGAATATGAC 960
 DB 903 TTGATGAGATGCTGTATGCGCTGCGCGCTTTTATAGCCGAAACACAGCTGAATATGAC 962
 QY 961 GACAACATTCGACAAAGCGCTCTCCAGCTGAACCAAAATATACAAATTCGCAAGACAC 1020
 DB 963 GACAACATTCGACAAAGCGCTCTCCAGCTGAACCAAAATATACAAATTCGCAAGACAC 1022
 QY 1021 CTCCTGAGGCGCTCTCTGACAGAGACAGCAAGACGCGCTCGGCGCAAAAGATATATTC 1080
 DB 1023 CTCCTGAGGCGCTCTCTGACAGAGACAGCAAGACGCGCTCGGCGCAAAAGATATATTC 1082
 QY 1081 ATGAGATTAAGAGTATGCTCTCTCTCTTAACTTAAGTGGATGATCTATATATATAT 1140
 DB 1083 ATGAGATTAAGAGTATGCTCTCTCTCTTAACTTAAGTGGATGATCTATATATATAT 1142
 QY 1141 AAGATTACTCCCGCTTTTAAACCAAAATGTGAGTGGCGCCCAACAGCTACGGTATTTGAC 1200
 DB 1143 AAGATTACTCCCGCTTTTAAACCAAAATGTGAGTGGCGCCCAACAGCTACGGTATTTGAC 1202
 QY 1201 CCGAGTTCACGAAAGAGCTGTCCCAACTCCATTTGGCAAGTCCCTGACAGAGTCTCTC 1260
 DB 1203 CCGAGTTCACGAAAGAGCTGTCCCAACTCCATTTGGCAAGTCCCTGACAGAGTCTCTC 1262
 QY 1261 GTCACAGCAGCGCTCAAGGAAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCTGCTGCT 1320
 DB 1263 GTCACAGCAGCGCTCAAGGAAGCTGCGGAGGCTTCTTCTAGGCTTTTCTATGCTGCTGCT 1322
 QY 1321 ACGGACTCTTCTCTGAAACCTGTTAGGCTTGGTTTAAAGGATTTATGCTGCTGCT 1380
 DB 1323 ACGGACTCTTCTCTGAAACCTGTTAGGCTTGGTTTAAAGGATTTATGCTGCTGCT 1382
 QY 1381 CGAATGTTTATGTTAGTTCCTTCTGAGCGCTGAGTGCAGAGCATCTTACAAAGAGAA 1440
 DB 1383 CGAATGTTTATGTTAGTTCCTTCTGAGCGCTGAGTGCAGAGCATCTTACAAAGAGAA 1442
 QY 1441 TTTGCAATCTCTGGAAGCTTACCAATCTTATGACACACTGTTCTGCTGGAA-TTTTTTTGA 1499
 DB 1443 TTTGCAATCTCTGGAAGCTTACCAATCTTATGGAACACTGTTCTGCTGGAAAGCTTTTGA 1502
 QY 1500 AGAGCAGATCTCTCTCAGTGAAGTCTATGAGGTTTTCATTTT-ATTCTTCTTCAAGT 1558
 DB 1503 AGAGCAGATCTCTCTCAGTGAAGTCTATGAGGTTTTCATTTTTCATCTCTTCCCTCAAGCT 1562
 QY 1559 GGTGCTATCTCTGAAACGAGCGTTAGAGTGCAGCGCTTACAGCGAGCGAGGAGTTCGTTA 1618
 DB 1563 GGTGCTATCTCTGAAACGAGCGTTAGAGTGCAGCGCTTACAGCGAGCGAGGAGTTCGTTA 1622
 QY 1619 GAAGCGGAC-CTGTTCTTAAAGAGTCTCTGCAAGTCTGCTGGCTGAGTGAACAGAA 1677
 DB 1623 GAAGCGGAGCGTGTCTTAAAGAGTCTCTGCAAGTCTGCTGGCTGAGTGAACAGAA 1682
 QY 1678 TATTATGAATGTGCTCTTCTGAAAGATGTGTTAGTCTCAAAAGCTTTTCTCTATGCA 1737
 DB 1683 TATTATGAATGTGCTCTTCTGAAAGATGTGTTAGTCTCAAAAGCTTTTCTCTATGCA 1742
 QY 1738 GTGTTGAGTCTCTTATTTTCTGATATGCTGTGTAACCGCTGCTGAGTGTG 1797

[illegible]


```
xx PS Claim 1: SEQ ID NO 772: 65pp; English.
xx CC The invention relates to a combination which, comprises a number of
xx CC polynucleotide probes comprising a sequence selected from one of the 1490
xx CC sequences mentioned in the specification. The combination is useful as an
xx CC array element in a microarray for monitoring the expression of a number
xx CC of target polynucleotides. The microarray is particularly useful in the
xx CC diagnosis and treatment of cancer and immunopathology and neuropathology.
xx CC The microarray is useful in diagnostics and treatment regimens, drug
xx CC discovery and development, toxicological and carcinogenicity studies,
xx CC forensic and pharmacogenomics. The microarray is also useful for
xx CC monitoring progression of diseases and for developing sophisticated
xx CC profiles for the effects of currently available therapeutic drugs. The
xx CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
xx CC and genomic fragments and in research and diagnostic applications. The
xx CC array can detect changes in expression in a large number of genes coding
xx CC for different signaling pathway populations which can be used to diagnose
xx CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
xx CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
xx CC and Parkinson's disease. The present sequence represents a polynucleotide
xx CC probe of the invention.
xx CC Note: The sequence data for this patent did not form part of the printed
xx CC specification but was obtained in electronic format directly from USPTO
xx CC at seqdata.uspto.gov/sequence.html?DocID=06500938b1.
xx SQ Sequence 2311 BP; 604 A; 508 C; 506 G; 692 T; 1 other;

Query Match          95.5%; Score 2262.8; DB 25; Length 2311;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2300; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY 34 CGCGTGGTATGACGGTGAACAACTGAGCTCTCTAGGGCACCCCTCACTTACTCCAGGATG 93
DB 1 CGCGTGGTATGACGGTGAACAACTGAGCTCTCTAGGGCACCCCTCACTTACTCCAGGATG 60

QY 94 AGGGGCGATGGGCAATCTCATCGCTTTCATGAGCAGAGAGAGGATGGGTCTGAACGAC 153
DB 61 AGGGGCGATGGGCAATCTCATCGCTTTCATGAGCAGAGAGAGGATGGGTCTGAACGAC 120

QY 154 TTTATTCAGAAGATTGCCAATAACTCCCTATGCATGCAACACCCCTGAAAGTTTCAGTCCATC 213
DB 121 TTTATTCAGAAGATTGCCAATAACTCCCTATGCATGCAACACCCCTGAAAGTTTCAGTCCATC 180

QY 214 TTGAAGATCTCCCAACCTCAGAGGCTGAGCTTATGAATGCGAACCCCTTCTCTCCACCA 273
DB 181 TTGAAGATCTCCCAACCTCAGAGGCTGAGCTTATGAATGCGAACCCCTTCTCTCCACCA 240

QY 274 AGTCCTTCTCAGCAATCAACCTTGGCCCGCTGCTCCCAATCTCATGCTAAACCATCTGAC 333
DB 241 AGTCCTTCTCAGCAATCAACCTTGGCCCGCTGCTCCCAATCTCATGCTAAACCATCTGAC 300

QY 334 TTTCACTTCTTGAAGATGATCGGAAGGGCAGTTTTGGAAAGTTTCTTACGACAGAC 393
DB 301 TTTCACTTCTTGAAGATGATCGGAAGGGCAGTTTTGGAAAGTTTCTTACGACAGAC 360

QY 394 AAGCAGAGAAGATGTTCTATGCTCAGTCAAGCTTTTACAGAGAGAGCAATCTGAAAG 453
DB 361 AAGCAGAGAAGATGTTCTATGCTCAGTCAAGCTTTTACAGAGAGAGCAATCTGAAAG 420

QY 454 AAAGAGGAGAAGCATATTATTCGGAGCGGAATGTTCTGTTGAAGAAATGTGAAGCACCCCT 513
DB 421 AAAGAGGAGAAGCATATTATTCGGAGCGGAATGTTCTGTTGAAGAAATGTGAAGCACCCCT 480

QY 514 TTCCGTGGGGCTTCACTTCTTCCAGACTGCTGACAAATGTGACTTTGTCTCTAGAC 573
DB 481 TTCCGTGGGGCTTCACTTCTTCCAGACTGCTGACAAATGTGACTTTGTCTCTAGAC 540

QY 574 TACATTAATGTGGAGAGTTGTTCTTACCATCTCCAGAGGAGACGCTGCTCTCGAACA 633
DB 541 TACATTAATGTGGAGAGTTGTTCTTACCATCTCCAGAGGAGACGCTGCTCTCGAACA 600

QY 634 CGGGCTCGTTCTATGCTGCTGAATAGCCAGTGCCTTGGGCTACCTGCAATCTCACTGAAC 693
DB 601 CGGGCTCGTTCTATGCTGCTGAATAGCCAGTGCCTTGGGCTACCTGCAATCTCACTGAAC 660

QY 694 ATCGTTTATAGAGACTTAAACACAGAGAAATATTTTGTAGATTACACGGGACATATGTC 753
DB 661 ATCGTTTATAGAGACTTAAACACAGAGAAATATTTTGTAGATTACACGGGACATATGTC 720

QY 754 CTTACTGATTCGGAGCTCTGCAAGGAGAACTTGAACACACAGACACAACTCCACCTTTC 813
DB 721 CTTACTGATTCGGAGCTCTGCAAGGAGAACTTGAACACACAGACACAACTCCACCTTTC 780

QY 814 TGTGGCAGCGCGGAGTATCTCGCACTGAGGTGCTTCATAAGCAGCCCTTATGACAGACT 873
DB 781 TGTGGCAGCGCGGAGTATCTCGCACTGAGGTGCTTCATAAGCAGCCCTTATGACAGACT 840

QY 874 GTGAGCTGGTGGCTGGGAGCTCTCTGTATGAGAGGCTGTATGGGCTGCCGCTTTT 933
DB 841 GTGAGCTGGTGGCTGGGAGCTCTCTGTATGAGAGGCTGTATGGGCTGCCGCTTTT 900

QY 934 TATAGCCGAAACACAGCTGAAATGTACGACAACTTCTGAAACAGCCCTCTCCAGCTGAAA 993
DB 901 TATAGCCGAAACACAGCTGAAATGTACGACAACTTCTGAAACAGCCCTCTCCAGCTGAAA 960

QY 994 CCAATATTCAAAATTCGCAAGACACCTCTCGAGGGCTCTCTGCAAGAGGACAGAC 1053
DB 961 CCAATATTCAAAATTCGCAAGACACCTCTCGAGGGCTCTCTGCAAGAGGACAGAC 1020

QY 1054 AAGGGCTCGGGCCCAAGGATGCTTCATGAGAGATTAGAGTCAATGCTTCTTCTCTT 1113
DB 1021 AAGGGCTCGGGCCCAAGGATGCTTCATGAGAGATTAGAGTCAATGCTTCTTCTCTT 1080

QY 1114 ATTAAGCTGGGATGATCTCATTATAAGAAAGATTACTCCCGCTTTTAAACCCAAATGTGAGT 1173
DB 1081 ATTAAGCTGGGATGATCTCATTATAAGAAAGATTACTCCCGCTTTTAAACCCAAATGTGAGT 1140

QY 1174 GGGCCCAACAGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGCCCTGTCCCAATGCC 1233
DB 1141 GGGCCCAACAGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGCCCTGTCCCAATGCC 1200

QY 1234 ATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCGTCAGAGAGCTTCCGAGGCT 1293
DB 1201 ATTGGCAAGTCCCTGACAGGCTCTCTGTCACAGCCAGCGTCAGAGAGCTTCCGAGGCT 1260

QY 1294 TTCTAGGCTTTTCTATGCGCTTCCACGGACTCTTTCTCTGAACTCTTAAAGCTT 1353
DB 1261 TTCTAGGCTTTTCTATGCGCTTCCACGGACTCTTTCTCTGAACTCTTAAAGCTT 1320

QY 1354 GGTTTTAAAGGATTTTATGTTGTTTCCGAATGTTTGTAGTTAGCTTTTGTGGAGCGC 1413
DB 1321 GGTTTTAAAGGATTTTATGTTGTTTCCGAATGTTTGTAGTTAGCTTTTGTGGAGCGC 1380

QY 1414 CAGCTGACAGGACATCTTTACAAAGAGAAATTTGCACATCTCTGGAAGCTTAGCAATCTTATT 1473
DB 1381 CAGCTGACAGGACATCTTTACAAAGAGAAATTTGCACATCTCTGGAAGCTTAGCAATCTTATT 1440

QY 1474 GCACACTGTTCCCTGGAA-TTTTTTGAAGACACATCTCTCTCAGTGTAGCTTATGAGGTT 1532
DB 1441 GCACACTGTTCCCTGGAA-TTTTTTGAAGACACATCTCTCTCAGTGTAGCTTATGAGGTT 1500

QY 1533 TTTATTTTATCTTCTTCCCAAGTGTGTATCTCTGAAACGAGGCTTAGAGTCCGCG 1592
DB 1501 TTTATTTTATCTTCTTCCCAAGTGTGTATCTCTGAAACGAGGCTTAGAGTCCGCG 1560

QY 1593 CTTAGCGGAGGCGAGGAGTTTCGTTAGAAAGCGGAC-CTGTCTTCTTAAAAAGGTTCTCTGCG 1651
DB 1561 CTTAGCGGAGGCGAGGAGTTTCGTTAGAAAGCGGAGCTGTCTTCTTAAAAAGGTTCTCTGCG 1620

QY 1652 AGATCTGTCGGGCTGTATGACGAAATATTATGAATGTGCTTCTTCTGAG-AGATGT 1710
DB 1621 AGATCTGTCGGGCTGTATGACGAAATATTATGAATGTGCTTCTTCTGAG-AGATGT 1680

QY 1711 GTTACCTCCAAAGCTTTTCTTATCTGAGTGTTCAGTTCTTATTTTCTCCCTGAGATAT 1770
DB 1711 GTTACCTCCAAAGCTTTTCTTATCTGAGTGTTCAGTTCTTATTTTCTCCCTGAGATAT 1770
```

Db 1681 GTTACCTCAAGCTTTTCCTATCCAGTGTTCAGTCTCTTATTTCCCTTGTGATAT 1740
 QY 1771 GCTGTGTAACCGTCTGTGAGTGTGATGCTGATCAGATGGATTTTGTATAAGC 1830
 Db 1741 GCTGTGTAACCGTCTGTGAGTGTGATGCTGATCAGATGGATTTTGTATAAGC 1800
 QY 1831 ATCAATGTGACACTTGCAGGACACTACAACGCTGGACATTTGTTGTTCTTCCATATTG 1890
 Db 1801 ATCAATGTGACACTTGCAGGACACTACAACGCTGGACATTTGTTGTTCTTCCATATTG 1860
 QY 1891 GAAGATAAATTTATGTGAGACTTTTGTGAAGTACGGTTAATACTAAATTTATTGA 1950
 Db 1861 GAAGATAAATTTATGTGAGACTTTTGTGAAGTACGGTTAATACTAAATTTATTGA 1920
 QY 1951 AATGTCCTTGCATGCTGATTCAGATGCTTAAGAAAGCATTTGCTGTACAATATT 2010
 Db 1921 AATGTCCTTGCATGCTGATTCAGATGCTTAAGAAAGCATTTGCTGTACAATATT 1980
 QY 2011 TCTATTTTGAAGAGGTTTTTATGACCAATGCCAGTTGTCAGTCAGAGCGTTGGT 2070
 Db 1981 TCTATTTTGAAGAGGTTTTTATGACCAATGCCAGTTGTCAGTCAGAGCGTTGGT 2040
 QY 2071 GTTTTTCATGTTTAAATGTCACCTGTAATGCGCATTTATTTATGTTTTTTTTTTC 2130
 Db 2041 GTTTTTCATGTTTAAATGTCACCTGTAATGCGCATTTATTTATGTTTTTTTTTTC 2100
 QY 2131 ATTCCTGATAATGTATGATGATAAAGACGCTCTGTACATTTGGGTTATAACACTAGT 2190
 Db 2101 ATTCCTGATAATGTATGATGATAAAGACGCTCTGTACATTTGGGTTATAACACTAGT 2160
 QY 2191 ATATTAAACTTACAGGCTTATTTGTAATGTAAGAACCCATTTTAAATGTAAGTAAT 2250
 Db 2161 ATATTAAACTTACAGGCTTATTTGTAATGTAAGAACCCATTTTAAATGTAAGTAAT 2220
 QY 2251 CATGGTTAATACGTACAACTCTTCCTCATCCATCCATCACACAACCTTTTGTGCTGA 2310
 Db 2221 CATGGTTAATACGNCACTCTTCCTCATCCATCCATCACACAACCTTTTGTGCTGA 2280
 QY 2311 TAACTGATTTGGTTGCAATAAACCCTG 2341
 Db 2281 TAACTGATTTGGTTGCAATAAACCCTG 2311

RESULT 9
 AA233647
 ID AA233647 standard; cDNA; 2281 BP.
 AC AA233647:
 DT 08-DEC-1999 (first entry)
 XX Human breast tumour-associated EST 37.
 DE Expressed sequence tag; EST; human; breast; cancer; gene therapy;
 XX treatment; tumour; cytostatic; medicament; ss.
 XX Homo sapiens.
 XX DE19813839-A1.
 XX 23-SEP-1999.
 XX 20-MAR-1998; 98DE-1013839.
 XX 20-MAR-1998; 98DE-1013839.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Specht T, Hinemann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX WPI; 1999-528981/45.
 XX P-PSDB; AA48573.

PT Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy -
 PS Claim 1a; 116-117; 188pp; German.
 XX This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AA233611-248617 represents expressed
 CC sequence tags described in the method of the invention.
 XX Sequence 2281 BP; 601 A; 498 C; 494 G; 688 T; 0 other;
 SQ

Query Match 90.3%; Score 2140.8; DB 20; Length 2281;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 2167; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 195 CCTGAAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGGAGCGCTTATGATTC 254
 Db 95 CCTGAAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGGAGCGCTTATGATTC 154
 QY 255 CAACCTCTCTCTCCACCAAGTCTTCTCAGCAATCAACCTTGGCCCGTGGTTC 314
 Db 155 CAACCTCTCTCTCCACCAAGTCTTCTCAGCAATCAACCTTGGCCCGTGGTTC 214
 QY 315 TCATGCTTAACCATCTGACTTTCATCTTCTGAAAGTATCGGAAGGCGACTTTTCAAA 374
 Db 215 TCATGCTTAACCATCTGACTTTCATCTTCTGAAAGTATCGGAAGGCGACTTTTCAAA 274
 QY 375 GGTCTCTCTAGCAAGACACCAAGGCAAGAAAGTGTCTTATGCACTCAAGTTCATAGAA 434
 Db 275 GGTCTCTCTAGCAAGACACCAAGGCAAGAAAGTGTCTTATGCACTCAAGTTCATAGAA 334
 QY 435 GAAGCAATCTCTGAAGAAAGAGAGAGAGATATTTATGTCGAGCGGATGTCGTT 494
 Db 335 GAAGCAATCTCTGAAGAAAGAGAGAGAGATATTTATGTCGAGCGGATGTCGTT 394
 QY 495 GAAGCAATCTCTGAAGAAAGAGAGAGAGATATTTATGTCGAGCGGATGTCGTT 554
 Db 395 GAAGCAATCTCTGAAGAAAGAGAGAGAGATATTTATGTCGAGCGGATGTCGTT 454
 QY 555 ATGTGACTTGTCTAGACTACATTAATGCTGGAGAGTGTCTTACCATCTCCAGAGGA 614
 Db 455 ATGTGACTTGTCTAGACTACATTAATGCTGGAGAGTGTCTTACCATCTCCAGAGGA 514
 QY 615 ACGCTGCTTCTCTGGAACCAAGGCTCTGTTTCTATGCTGCTGAAATAGCCAGTGTGG 674
 Db 515 ACGCTGCTTCTCTGGAACCAAGGCTCTGTTTCTATGCTGCTGAAATAGCCAGTGTGG 574
 QY 675 CTACCTGCAATTCACATCAACATCGTTTATAGAGACTTAAACACAGAGATATTTGCTAGA 734
 Db 575 CTACCTGCAATTCACATCAACATCGTTTATAGAGACTTAAACACAGAGATATTTGCTAGA 634
 QY 735 TTCACAGGAGACATTTGCTTACATGCTGCTGAGTCTGCAAGGAGAGACATTCATACAA 794
 Db 635 TTCACAGGAGACATTTGCTTACATGCTGCTGAGTCTGCAAGGAGAGACATTCATACAA 694
 QY 795 CAGCACAACATCCACCTCTGTGGCAGCCGCGAGTATCTCCACCTCAGGTTGCTTCATA 854
 Db 695 CAGCACAACATCCACCTCTGTGGCAGCCGCGAGTATCTCCACCTCAGGTTGCTTCATA 754
 QY 855 CGAGCTTATGACAGGACTGTGGAGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 914
 Db 755 CGAGCTTATGACAGGACTGTGGAGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 814
 QY 915 GTATGGCTTGTGGCTTTTATGACGCAACACAGCTGGAATGTACGACACATTCGAA 974
 Db 815 GTATGGCTTGTGGCTTTTATGACGCAACACAGCTGGAATGTACGACACATTCGAA 874
 QY 975 CAAGCTCTCCAGCTGAAACCAATATTTACAAATTTCCGCAAGACACCTCTCTGGAGGCT 1034

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 SQ Sequence 2432 BP; 586 A; 586 C; 560 G; 699 T; 1 other;

Query Match	63.0%;	Score 1492.4;	DB 22;	Length 2432;
Best local Similarity	82.0%;	Pred. No. 0;		
Matches 1970;	Conservative	0;	Mismatches 332;	Indels 99; Gaps 18;
QY 23	TCTGTCTCCGCGTGTGTGATGACGGTGAATACTCTCATCGCTTTCATGAGGAGGAGGATGG 142			
DB 55	TCTTTCCGGGATCGGAGCGATGACCGTCAAGGCCGAGGCTGCTCGAAGCACCCCTACCT 114			
QY 83	ACTCAGGATGAGGGGATGCTGGCAATCTCTCATCGCTTTCATGAGGAGGAGGATGG 174			
DB 115	ACTCAGGATGAGGGGATGCTGGCAATCTCTCATCGCTTTCATGAGGAGGAGGATGG 174			
QY 143	GTCAGGAGCTTTATTCAGAGATGTCGAATACTCTCATGAGGAGGAGGATGG 202			
DB 175	GCTTCAAGGATTTATTCAGAGATGTCGAATACTCTCATGAGGAGGAGGATGG 234			
QY 203	TTCAGTCCATCTGAGGATCTCCCAACTCTCAGGAGGCTGAGCTTATGAATGCCACCCCT 262			
DB 235	TTCAGTCCATCTGAGGATCTCCCAACTCTCAGGAGGCTGAGCTTATGAATGCCACCCCT 262			
QY 263	CTCCTCCAGGATCTTCTCAGCAATCAACCTTGGCCGCTGCTCCCAATCTCTCATGCTA 322			
DB 295	CTCCTCCAGGATCTTCTCAGCAATCAACCTTGGCCGCTGCTCCCAATCTCTCATGCTA 322			
QY 323	AACTCTGAGCTTCTCAGTCTTGAAGGATGTCGAAGGAGGAGGATGG 382			
DB 355	AACTCTGAGCTTCTCAGTCTTGAAGGATGTCGAAGGAGGAGGATGG 414			
QY 383	TAGCAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 442			
DB 415	TAGCAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 474			
QY 443	TCTGAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 502			
DB 475	TCTGAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 534			
QY 503	TGAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 562			
DB 535	TGAGGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 594			
QY 563	TGCTCTGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 622			
DB 595	TGCTCTGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 654			
QY 623	TGCTGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 682			
DB 655	TGCTGAGGAGGAGGAGGATGTCGAAGGAGGAGGATGG 714			
QY 683	ATTCAGTCAACATCTGTTATAGAGCTTAAACAGGAGGATGTCGAAGGAGGAGGATGG 742			
DB 715	ATTCAGTCAACATCTGTTATAGAGCTTAAACAGGAGGATGTCGAAGGAGGAGGATGG 774			
QY 743	GACAGGATGCTTCTGAGGAGGATGTCGAAGGAGGAGGATGG 802			
DB 775	GACAGGATGCTTCTGAGGAGGATGTCGAAGGAGGAGGATGG 831			
QY 803	CATCCAGGATGCTGTCGAGGAGGATGTCGAAGGAGGAGGATGG 862			
DB 832	CATCCAGGATGCTGTCGAGGAGGATGTCGAAGGAGGAGGATGG 891			
QY 863	ATGACAGGAGGATGTCGTCGAGGAGGATGTCGAAGGAGGAGGATGG 922			
DB 892	ATGACAGGAGGATGTCGTCGAGGAGGATGTCGAAGGAGGAGGATGG 951			

QY 923	TGCGGCTTTTATAGCCGAAACACAGCTGAATGTACGACACATTTCTGACAAATCTC 982			
DB 952	TGCGGCTTTTATAGCCGAAACACAGCTGAATGTACGACACATTTCTGACAAATCTC 1011			
QY 983	TCCAGCTGAAACCAATATTTACAAATTTCCGCAAGACACCTCTCGGAGGCGCTCTGACGA 1042			
DB 1012	TCCAGCTGAAACCAATATTTACAAATTTCCGCAAGACACCTCTCGGAGGCGCTCTGACGA 1071			
QY 1043	AGGACAGGACAAAGCGCTCGGCTCAAGGATGACTTCATGGAGATTTAGAGATTTAGTCT 1102			
DB 1072	AGGACAGGACAAAGCGCTCGGCTCAAGGATGACTTCATGGAGATTTAGAGATTTAGTCT 1131			
QY 1103	TCTTCTCTTAAATTAACGGGATGATCTCATTAATAAGAGATTTACCTCTCTTTAAATC 1162			
DB 1132	TCTTCTCTTAAATTAACGGGATGATCTCATTAATAAGAGATTTACCTCTCTTTAAATC 1191			
QY 1163	CAATGTCTGAGTGGGCGCAACAGCTTACGGCCTTTGACCCGAGCTTTACCGAAGATTTG 1222			
DB 1192	CAATGTCTGAGTGGGCGCAACAGCTTACGGCCTTTGACCCGAGCTTTACCGAAGATTTG 1251			
QY 1223	TCCCAACTCATTTGCGCAAGTCCCTGACAGCTTCTCTGTCACAGGAGGCTTAAAGAG 1282			
DB 1252	TCCCAACTCATTTGCGCAAGTCCCTGACAGCTTCTCTGTCACAGGAGGCTTAAAGAG 1311			
QY 1283	CTGCGGAGCTTCTCTAGGCTTTTCTTATGCGCTCCACGAGCTTCTTCTCTTAAAGCT 1342			
DB 1312	CAGCAGAGGCTTCTCTAGGCTTTTCTTATGCGCTCCACGAGCTTCTTCTCTTAAAGCT 1371			
QY 1343	TGTTAGGCTTGGTTTTAAAGGATTTTATGCTGTTTTCGGAATGT-TTTAGTTAGCTTT 1401			
DB 1372	--TCCCGGGATGGTCTGAGGAGCTTCTCAGCGCTTCTCTTAAAGTGTTCGTTAGGCT 1429			
QY 1402	TGCTGGAGCGGCGGCTGACAGGAGCTTCTTAAAGAGATTTGACATCTCTGAGAGCTT 1461			
DB 1430	TGCTGGAGCTTCCAGCTGACAGAGATTTTAAAGAGATTTGACACACACACACACAC 1485			
QY 1462	AGCAATCT-----TATGACACACTGTCTGCTGGAATTTTTCG 1498			
DB 1486	GCACTCTCGGCTGCGCGGCTGCGGCGGCTGCTGCTGATGGAGTCTTCT 1545			
QY 1499	---AAGAGCACATCTCTCTGAGCTGAGCTTCTTCAATTTTATTTCTCTCTCTCTCTCT 1555			
DB 1546	GAGAGCACACCTCTCTCTCAATGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1605			
QY 1556	CTGCTGCT 1612			
DB 1606	CTGCTGCT 1665			
QY 1613	TC-GTTAGAAAGCGGAGCTTCTTCTTAAAGAGCTTCTCTGAGAGATCTCTCTGCTGCTGAT 1671			
DB 1666	TCAGTTAGAGAGAGATGACAGCTCTAGAGGAATCCCGGAGGCTCTCTCTGAGCTTCTGAT 1725			
QY 1672	GACAGATATTAAGATGCTCTTCTTCTGAGAGATTTGCTGAGCTCTCTCTCTCTCTCTCTCT 1730			
DB 1726	CAGAGATATTTCTGCAATGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1782			
QY 1731	TATCGCAGCT-----TTCAGTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1764			
DB 1783	TATCGCAGCT 1842			
QY 1765	GATATGCTGCTGAGAGGCT 1824			
DB 1843	GATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1901			
QY 1825	ATAAGCATCAATGTCACACTTTCAGGACACTACAGCTGGGAGATTTGTTTCTCTCTCTCT 1884			
DB 1902	GTGAGCTCAATGTCACACTTTCAGGACACTACAGCTGGGAGATTTGTTTCTCTCTCTCTCT 1961			
QY 1885	TATTTGGAGATCAATTTATGCTGAGACTTTTCTGAGATPACGGTGAATTAACATAAAT 1944			
DB 1962	CATTTGGAGATCAATTTATGCTGAGACTTTTCTGAGATPACGGTGAATTAACATAAAT 2021			

RESULT 11	
AAA46592	
ID	AAA46592 standard; DNA; 2365 BP.
XX	
XX	AAA46592;
XX	
XX	
XX	25-SEP-2000 (first entry)
XX	
DE	DNA encoding a rat serum and glucocorticoid induced protein kinase.
XX	
XX	Protein kinase; Pkh1; Pkh2; Ypk1; Yrk2; protein kinase B-alpha;
KW	serum and glucocorticoid induced protein kinase; SGK; PKAalpha;
KW	3-phosphoinositide-dependent protein kinase-1; PDK1; fungal infection;
KW	thrush; cancer; diabetes; obesity; antifungal; Candida infection; ss.
XX	
OS	Rattus sp.
XX	
Key	Location/Qualifiers
FT	1..1288
CDS	/*tag= a
FT	/product= "serum and glucocorticoid induced protein
FT	kinase (SGK)"
FT	/transl_except= (pos: 1, aa: Thr)
FT	/note= "the codon encoding Met at position 1 and
FT	Pro at position 318 are not given"
XX	
XX	WO200036135-A2.
XX	
XX	
PD	22-JUN-2000.
XX	
XX	14-DEC-1999; 99WO-GB04228.
PF	
XX	
XX	14-DEC-1998; 98US-0112114.
PR	
XX	
XX	(MEDI-) MEDICAL RES COUNCIL.
PA	(REGC) UNIV CALIFORNIA.
PA	
XX	
XX	Thorner JW, Alessi DR, Torrance PD, Casamayor A;
PI	
XX	WPI; 2000-442391/38.
XX	
DR	

QY	109	ATTCTCATCGCTTTCTATGAAAGCAGGAGATGGGTCTGAACGACTTTATTTCAGAAATTT	168
Db	62	ATCTCTCATCGCTTTCTATGAAACAGAGAAAGATGGGCTGAAGCATTTTATTTCAGAAATTT	121
QY	169	GCCATAATCTCTATGCAATGCAAAACCCCTCAAGTTTCAGTCCATCTTGAAGATATCCCAA	228
Db	122	GCCAACTCTATGCAATGCAAAACCCCTGAAGTTCAATCTATTGAAATATCTCCCAA	181
QY	229	CCTCAGGAGCCTGAGCTTATGAATGCCAACCCCTTCCTCCACCAAGTCTTCTTCAGAA	288
Db	182	CCTCAGGAGCGCAACTTATGAACGCCAACCCCTCACCCTCCCAAGTCCCTTCACAA	241
QY	289	ATCAACCTTGCCCGCTGCTCAATCCTCATGCTAAACCATCTGACTTTCACATCTTCGAAA	348
Db	242	ATCAACCTTGGGTCCATCTCMAATCCCGACGCCAACCCCTCTGACTTCCTCATCTTCGAAA	301
QY	349	GTGATCGGAAGGCGAGTTTGGAAAGGTTCTTCTACCAAGACACAAAGCAGAAAGTGT	408
Db	302	GTGATCGGAAGGCGAGTTTGGAAAGGTTCTTCTACCAAGGACAAAGCAGAAAGCA	361
QY	409	TTCATGCACTCAAGTTTACAGAAAGAACCAATCCTGNAAGAAAGACAGCAGACAT	468
Db	362	TTCATGCGGTCAAAAGTTTGCAGAAAGAACCCATCTTGAAGAAAGGAGGAGGAGCAT	421
QY	469	ATTATGTCGGAGCGGAATGTCTCTTGAAGAATGTGAAGCACCCCTTTCCTGGTGGGCGCT	528
Db	422	ATTATGTCAGAGCGCAATGTCTCTTGAAGAATGTGAAGCACCCCTTTCCTGGTGGGCGCT	481
QY	529	CACCTCTCTTTCCAGACTGCTGACAAATGTTACTTTGTCGTAGACTACATTAATGCGTGA	588
Db	482	CACCTCTCTTTCCAGACTGCTGACAAATGTTACTTTGTCGTAGACTACATTAATGCGGGA	541
QY	589	GAGTTGTCTTACCATCTCCAGAGGAAGCTGCTTCTCGGAACCCAGGCGCTGTTCTTCTAT	648
Db	542	GAGCTGTCTTACCATCTCCAGAGGAAGCTGCTTCTCGGAACCCAGGCTGCTTCTTCTAT	601
QY	649	GCTGCTGAATAGCCAGTGGCTTGGCGCTTACCTGCATCTCACTGAACATGCTTTATAGAGC	708
Db	602	CGAGCTGAATAGCCAGTGGCTTGGCTTATCTGCATCTCCCTAAACATGTTATCGAGC	661

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;
 XX WP1: 2001-032161/04.
 DR P-PSDB; AAB65613.
 XX
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers -
 XX
 XX Disclosure: Fig 2; 310pp; English.
 XX
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure,
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and
 CC reproductive disorders.
 XX
 XX Sequence 1296 BP; 345 A; 333 C; 293 G; 325 T; 0 other;
 SQ

Query Match 54.7%; Score 1296; DB 22; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 7.1e-281;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

43 ATGACGGTGAACACTGAGGCTGCTTAAGGCGACCTCACTTACTCCAGGATCAGGGGATG 102
 1 ATGACGGTGAACACTGAGGCTGCTTAAGGCGACCTCACTTACTCCAGGATCAGGGGATG 60
 103 GTGGCAATTTCTCATCGCTTTTATGAGCAGAGGAGGATGGTCTGACAGCACTTTATTTCAG 162
 61 GTGGCAATTTCTCATCGCTTTTATGAGCAGAGGAGGATGGTCTGACAGCACTTTATTTCAG 120
 163 AGATTGGCAATTAACCTCTATGATGCAACACCTGAAATTCAGTTCATCTCAAGATC 222
 121 AGATTGGCAATTAACCTCTATGATGCAACACCTGAAATTCAGTTCATCTCAAGATC 180
 223 TCCCAACCTCAGAGGCGCTGAGCTTATGAAATGCCAACCTTCTCCCAACCAAGTCTTCT 282
 181 TCCCAACCTCAGAGGCGCTGAGCTTATGAAATGCCAACCTTCTCCCAACCAAGTCTTCT 240
 283 CAGCAATCAACCTTGGCGCGTCTGCTCAATCTCATGCTAAACCATCTGACTTTCACCTTC 342
 241 CAGCAATCAACCTTGGCGCGTCTGCTCAATCTCATGCTAAACCATCTGACTTTCACCTTC 300
 343 TTGAAAGTATCGGAAAGGCGGAGCTTTTGAAGGTTCTTCTAGCAAGACACAAAGCGAA 402
 301 TTGAAAGTATCGGAAAGGCGGAGCTTTTGAAGGTTCTTCTAGCAAGACACAAAGCGAA 360
 403 GAAGTCTTATGCTGAGTGAAGTCTTATGAGGAGGAGCAATCTGAAAGAGAGAGAG 462
 361 GAAGTCTTATGCTGAGTGAAGTCTTATGAGGAGGAGCAATCTGAAAGAGAGAGAG 420
 463 AAGCATATTATGCTGAGGAGGAGTCTTCTGTTGAAGATGTAAGACACCTTCTCGGTG 522
 421 AAGCATATTATGCTGAGGAGGAGTCTTCTGTTGAAGATGTAAGACACCTTCTCGGTG 480
 523 GGCTTTCACCTTCTTTCAGACTCTGACAAATTTGACTTTGCTTACAGTACTATTAAT 582
 481 GGCTTTCACCTTCTTTCAGACTCTGACAAATTTGACTTTGCTTACAGTACTATTAAT 540
 583 GGTGGAGAGTGTGTTTCTTACCATCTCCAGAGGAGCAACCTGCTTCTGGAACCAAGGCGCTG 642

541 GGTGGAGAGTGTGTTTCTTACATCTCCAGAGGAGAGCGTCTTCTGCGAACCACGAGTCT 600
 643 TTCTATGCTGCTGAATAGCAGTCCCTTGGGCTTACCTGCATTCACCTGACATGCTTAT 702
 601 TTCTATGCTGCTGAATAGCAGTCCCTTGGGCTTACCTGCATTCACCTGACATGCTTAT 660
 703 AGAGACTTAAACACAGAGATATTTTCTAGATTTCAGGAGACACATTTGCTTACTGAT 762
 661 AGAGACTTAAACACAGAGATATTTTCTAGATTTCAGGAGACACATTTGCTTACTGAT 720
 763 TTGCGACTCTCCAAAGGAGAACATTGAACACACAGCACAACATCCACCTTCTCTGTCAG 822
 721 TTGCGACTCTCCAAAGGAGAACATTGAACACACAGCACAACATCCACCTTCTCTGTCAG 780
 823 CCGGAGTATCTCGACCTCGAGTCTTCTAAGCAGCTTATGACAGGAGCTGTCGATG 882
 781 CCGGAGTATCTCGACCTCGAGTCTTCTAAGCAGCTTATGACAGGAGCTGTCGATG 840
 883 TGGTCCCTGGAGCTGTCTTATGAGATGCTATGCGCTGCGCTCTTTTATATATGA 942
 841 TGGTCCCTGGAGCTGTCTTATGAGATGCTATGCGCTGCGCTCTTTTATATATGA 900
 943 AACACAGCTGAATATGACACACATTTTGAACAGCTCTCTAGCTGACATGAATAT 1002
 901 AACACAGCTGAATATGACACACATTTTGAACAGCTCTCTAGCTGACATGAATAT 960
 1003 ACAAAATTCGCAAGACACCTCTCTGAGGCGCTCTGAGAGGAGCAGACACAAAGCTC 1062
 961 ACAAAATTCGCAAGACACCTCTCTGAGGCGCTCTGAGAGGAGCAGACACAAAGCTC 1020
 1063 GGGGCCAAGGATGACTTCTATGAGATTAAGAGTCAATGTCTTCTCTTATATATGA 1122
 1021 GGGGCCAAGGATGACTTCTATGAGATTAAGAGTCAATGTCTTCTCTTATATATGA 1080
 1123 GATGATCTCATTAATAAGAGATTAATCTCCCTCTTTAACCCAAATGTGAGTGGGCAAC 1182
 1081 GATGATCTCATTAATAAGAGATTAATCTCCCTCTTTAACCCAAATGTGAGTGGGCAAC 1140
 1183 GAGCTACGGCAGCTTTGACCCGAGTTTACCAAGAGCTGTCCCACTCCATTCGCAAG 1242
 1141 GAGCTACGGCAGCTTTGACCCGAGTTTACCAAGAGCTGTCCCACTCCATTCGCAAG 1200
 1243 TCCCTGTAGCAGCTCTGCTACAGCTGAGCTCAAGAGAGCTGCGGAGTCTTCTATGAG 1302
 1201 TCCCTGTAGCAGCTCTGCTACAGCTGAGCTCAAGAGAGCTGCGGAGTCTTCTATGAG 1260
 1303 TTTTCTATGAGCTCTGCTACAGCTGAGCTCAAGAGAGCTGCGGAGTCTTCTATGAG 1338
 1261 TTTTCTATGAGCTCTGCTACAGCTGAGCTCAAGAGAGCTGCGGAGTCTTCTATGAG 1296

RESULT 13

AAS44987

ID AAS44987 standard; cDNA: 1366 bp.

XX AAS44987;

XX AC

XX DT

XX 18-DEC-2001 (first entry)

XX DE

XX cDNA encoding novel human secretory protein, Seq ID No 58.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; parkinson's disease; burn;
 KW ankyrotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen; ss.

XX OS
 XX Homo sapiens.

XX PN W0200166689-A2.
XX PD 13-SEP-2001.
XX PF 05-MAR-2001; 2001WO-US04942.
XX PR 07-MAR-2000; 2000US-0519705.
XX PR 19-MAY-2000; 2000US-0574454.
XX PR 17-JUN-2000; 2000US-0596193.
XX PR 14-JUL-2000; 2000US-0616847.
XX PR 19-SEP-2000; 2000US-0665363.
XX PR 20-OCT-2000; 2000US-0693267.
XX PA (HYSSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
XX PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
XX DR P-PSDB; AA028087.
XX WIPI: 2001-589934/66.
XX DR Novel polypeptides and nucleic acids obtained from cDNA libraries
XX PT prepared from various human tissues, for diagnosis and treatment of
XX PT cancer, neurological, inflammatory, and autoimmune disorders.
XX PS Claim 1; SEQ ID No 68; 107pp; English.
XX CC The invention relates to novel isolated human secreted polypeptides (I)
XX CC and polynucleotides (II). (I) and (II) are useful for treating
XX CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
XX CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
XX CC involved in increasing haematopoiesis, stem cell survival, bone growth
XX CC and remodeling. (I), (II) and modulators of (II) are useful for
XX CC prophylaxis or treatment of one or more cancers. (II) is also useful for
XX CC creating transgenic animals useful for studying the in vivo activities of
XX CC the polypeptide as well as for studying modulators of the polypeptides.
XX CC (I) induces the proliferation of neural cells and regeneration of nerve
XX CC and brain tissue and is useful for the treatment of central and
XX CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
XX CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
XX CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
XX CC activity, regulation of haematopoiesis and is useful for treating myeloid
XX CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
XX CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
XX CC tissue growth, and in tissue repair, healing of burns, incisions,
XX CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
XX CC disorders, or periodontal disease. Furthermore, (I) is also useful for
XX CC gut protection or regeneration and treatment of lung or liver fibrosis,
XX CC reperfusion injury in various tissues, various immune deficiencies and
XX CC disorders including severe combined immunodeficiency (SCID), bacterial or
XX CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
XX CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
XX CC reactions and conditions, such as asthma, or other respiratory problems.
XX CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
XX CC fertility, metabolism, catabolism, anabolism, storage or elimination of
XX CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
XX CC analgesic effects or other pain reducing effects, immunoglobulin like
XX CC activity and can act as an antigen in a vaccine composition to raise an
XX CC immune response. AAS44920-AAS45295 represent novel human secreted protein
XX CC coding sequences of the invention.
XX XX Sequence 1366 BP; 331 A; 374 C; 377 G; 284 T; 0 other;
XX Query Match 22.2%; Score 527.2; DB 22; Length 1366;
XX Best Local Similarity 69.5%; Pred. NO. 2.2e-108;
XX Matches 734; Conservative 0; Mismatches 313; Indels 9; Gaps 1;
XX QY 261 TTCTCTCCACCAAGTCTCTTCTACGACCAATCACTTGGCCGCTGCTCCATCTATGC 320
XX Db 213 TCCACAGCCCTCCAGGCCAATGGGAACATCACTTGGCCGCTTCAACCAACCAATGC 272
XX QY 321 TAAACATCTGACTTTTCACTTCTTGAAGATGATCGAAAGGCGAGTTTGGAAAGGTTCT 380

Db 273 CCAGCCACGGGACTTGGACTTCTCAAGTCTCATCGCAAGGGAACACTACGGGAAGTCTCT 332
QY 381 TCTAGCAAGACACCAAGCCAGCAAGAGTCTCTATGCAAGTCAAGTCTTTTACAGCAAGAAC 440
Db 333 ACTGGCAAGCGCAAGTCTCTATGGGCGTTCTATGCAAGTGAAGGTACTTACAGCAAGAAC 392
QY 441 AATCTGAAAAAAGAAAGAGAGAGATATTTATGTCGGAGCGGAATTTCTTTCTTGAAGAA 500
Db 393 CATCTTAAAGAAAGAAAGAGAGAGATATTTATGTCGGAGCGGAATTTCTTTCTTGAAGAA 452
QY 501 TGTGAAGCAAGCCCTTCTGTCGGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 560
Db 453 CGTGGCGACCCCTTCTGTCGGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 512
QY 561 CTCTTCTCTAGACTACATTAATGCTGAGAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTT 620
Db 513 CTCTGCTCTGACTATCTCAACGGGGAGAGACTCTTCTTCTTCTTCTTCTTCTTCTTCTT 572
QY 621 CTCTCTGGAACACCGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 680
Db 573 GTTCTCTGAGCCCGCCGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 632
QY 681 GCATCTCACTGAACATCTGTTTATAGAGACTTAAACACAGAGATATTTTCTTCTTCTTCT 740
Db 633 GCATCTCTCAACATCTTACAGGATCTGAAACACAGAGACATCTTCTTCTTCTTCTTCT 692
QY 741 GGGACACATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 800
Db 693 GGGACACATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 752
QY 801 AACATCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 860
Db 753 CACATCTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 812
QY 861 TTATGACAGGACTGTGAGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 920
Db 813 TTATGATGAGACAGTGGAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 872
QY 921 CTTCTGCGCTTTTATAGCGGAAACACAGCTGAAATGTAGCAACATCTTCTTCTTCTTCT 980
Db 873 CTTCTGCGCTTTTATAGCGGAAACACAGCTGAAATGTAGCAACATCTTCTTCTTCTTCT 932
QY 981 TCTCCAGCTGAACCAATATTTACAAATTTGCGAGACACTCTCTTCTTCTTCTTCTTCT 1040
Db 933 GCTACAGATCCCCGGAGGCGGACAGCTGCGGCGCTTCTTCTTCTTCTTCTTCTTCTTCT 992
QY 1041 GAAGACAGGACAAAGCGGCTGCGGCGGCAAGGATGACTTCTTCTTCTTCTTCTTCTTCT 1100
Db 993 CAAGACAGGAGGAGCGGCTGCGGCTGCGGCGGCAAGGATGACTTCTTCTTCTTCTTCT 1052
QY 1101 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1160
Db 1053 ATCTTCTGAGCCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1112
QY 1161 CCCAAATGTAGTGGGCGGCAAGGATGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1220
Db 1113 CCCAAATGTAGGAGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1172
QY 1221 TGTCTCCCACTCAATTTGCAAGTCCCTGAGAGGCTCTTCTTCTTCTTCTTCTTCTTCT 1280
Db 1173 TGTCTCCCACTCAATTTGCAAGTCCCTGAGAGGCTCTTCTTCTTCTTCTTCTTCTTCT 1223
QY 1281 AGCTGCGGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1316
Db 1224 GGCCTCAAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1259
XX RESULT 14
XX AAF44737
XX ID AAF44737 standard; cDNA: 1812 BP.
XX AC AAF44737;

OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 88..1191
 FT /tag= a
 FT /product= "human SGK2-alpha protein"
 XX WO200224947-A2.
 PN 28-MAR-2002.
 PD 20-SEP-2001; 2001WO-1B02237.
 PF 20-SEP-2000; 2000US-233999P.
 PR 02-OCT-2000; 2000US-237419P.
 PR 02-OCT-2000; 2000US-237423P.
 PR 04-OCT-2000; 2000US-238558P.
 PR 10-MAY-2001; 2001US-290555P.
 XX (KINE-) KINETEK PHARM INC.
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 XX Yoganathan T, Delaney AD;
 PI WPI; 2002-394145/42.
 DR P-PSDB; AAE22765.
 XX Diagnosing cancer, comprises determining the upregulation of expression
 of a nucleic acid sequence encoding a protein kinase or upregulation of
 expression of the protein kinase, in the cancer
 XX Claim 16; Page 65-66; 87pp; English.
 XX The invention relates to a method for screening biologically active agent
 that modulates cancer associated protein kinase function. The invention
 also relates to a method for diagnosing cancer comprising determining the
 upregulation of expression of a nucleic acid sequence encoding a protein
 kinase. The method is useful for diagnosing cancer. A protein kinase is
 useful for screening biological agents that modulate cancer associated
 protein kinase function. Downregulating the activity of protein kinase is
 useful for inhibiting the growth of a cancer cell, e.g. liver or colon
 cancer. A nucleic acid encoding protein kinase is useful to screen biopsy
 derived tumours and inflammatory samples such as arthritic synovium, for
 amplified DNA in the cell or increased expression of corresponding mRNA.
 or protein and is also useful to detect differences in expression levels
 such as molecular weight, amino acid and nucleotide sequences between the
 two cells. The present sequence is human serum and glucocorticoid-induced
 protein kinase, SGK2-alpha gene.
 XX Sequence 1812 BP; 461 A; 470 C; 470 G; 411 T; 0 other:
 Query Match 22.2%; Score 527.2; DB 24; Length 1812;
 Best Local Similarity 69.5%; Pred. No. 2.4e-108;
 Matches 734; Conservative 0; Mismatches 313; Indels 9; Gaps 1;
 QY 261 TTCTCTCCACCAAGTCCTTCACGAAATCAACCTTGGCCGCTGCTCAATCTCATGCG 320
 DB 117 TCCACAGCCCTCCAGGCCCAANTGGGAACATCAACCTGGGGCTTCAGGCCCAACCCAAATGC 176
 QY 321 TAAACCATCTCTTCACTTCTGAAAGTGATCGAAGGGCAGTTTGGAAAGGTTCT 380
 DB 177 CCAGCCACAGGCTTCGACTTCTCAAGTATCGCAAGGGAAGTACGGGAAGTCT 236
 QY 381 TCTAGCAAGACACAGGACAGAGAGTGTTCATGAGTCAGTCAAGTTCCTACAGAAAGC 440
 DB 237 ACTGGCCAAAGCGAAGTCTGATGGGGCTTCTATGCGAGTGAAGGTACTACAGAAAGTC 296
 QY 441 AATCCTGAAGAAAGAGAGAGAGATATATGTCGAGGCGGAATGTTCTGTTGAAGAA 500
 DB 297 CAGCTTAAGAAAGAGAGAGAGAGAGATATGTCGAGGCGGAATGTTCTGTTGAAGAA 356
 QY 501 TGTGAGACACCTTTCCTGCTGGGCTTTCACCTTCTTTCACAGACTGCTGACAAATGTA 560
 XX 111 111111 111111 111111 111111 111111 111111 111111 111111 111111

DB 357 CGTGGGCGACCCCTTCTCTGCGGCTTCCCTTCCAGACACCTGAGAGAGCTA 416
 QY 561 CTTTCTCTTAGACTACATTAAATGCTGGAGAGTCTTCTTACCATCTCCAGAGGAACTG 620
 DB 417 CTTCTGCTCGACTATGTCACAGCGGAGAGACTCTTCTTCCACCTGCGAGGAGAGAG 476
 QY 621 CTTCTGGAACACAGCGCTTCTTCTGCTGCTGAATATGACCTGCTGCTGCTGCTGCT 680
 DB 477 GTTCTCTGGAGCCCGGCGGAGGTTCTAGCTGCTGAGTGGCCAGCGCTAGCT 536
 QY 681 GCATTCACTGAACATCTTATAGAGACTTAAACCCAGAGAAATTTTGTACATTCACA 740
 DB 537 GCATCTCTCAACATCTTACGCGATCTGAACCCAGAGAGAACTTCTTGGACTGCCA 596
 QY 741 GGCACACATTTGCTTACTGATTTGCGACTCTGCAAGGAGAACATTCGACACACACAC 800
 DB 597 GGCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656
 QY 801 AACATCCACCTTCTGCTGCGAGCGGAGTATCTGCGACCTGAGTCTTCTATAGAGAG 860
 DB 657 CACATCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
 QY 861 TTATGACAGGACTGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 DB 717 TTATGATCGAGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
 QY 921 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
 DB 777 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
 QY 981 TCTCCAGCTGAACACCAATATTTACAAATTCGCAAGACACCTCTCTGAGGAGGCTGCTGCA 1040
 DB 837 GCTACAGATCCCGGAGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896
 QY 1041 GAAGGACAGGACAAAGCGCTGCGGCGCAAGGATGCTTCTGAGAGATTAAGAGCTGCT 1100
 DB 897 CAGGACACAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956
 QY 1101 CTTCTTCTCTTAAATTAAGTGGAGTCTCATTAAGAGATTAAGTCTGCTGCTGCTGCTGCT 1160
 DB 957 ATTCTTACGCCCAATAAAGTGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016
 QY 1161 CCCAAATGTGAGTGGCGCCCAAGGAGCTAGCGACTTTCACCCGAGTTTACCGAGAGCG 1220
 DB 1017 CCCAAATGTGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076
 QY 1221 TGTCCCAACTCCATTTGGCAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
 DB 1077 TGTGTCGAAGTCCATTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
 QY 1281 AGCTGCGGAGCTTCTTCTAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1316
 DB 1128 GGCTCAAGTGCATTCCTGCGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1163

Search completed: August 11, 2003, 02:31:22
 Job time : 624 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 11, 2003, 02:16:03 ; Search time 4491 Seconds
(without alignments)
12826.009 Million cell updates/sec

Title: US-10-000-039A-1

Perfect score: 2370

Sequence: 1 CACGAGGAGCGCTACGTC.....AAAAAAAAAAAAAAAAAAAA 2370

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EFT: *

1: em_estba:*

2: em_estba:*

3: em_estba:*

4: em_estba:*

5: em_estba:*

6: em_estba:*

7: em_estba:*

8: em_estba:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1501.8	63.4	2435	11 AK086892	AK086892 Mus muscu
2	1001.4	42.3	1059	13 BX386726	BX386726 BX386726
3	964.6	40.7	1052	13 BX386725	BX386725 BX386725
4	916.4	38.7	1072	12 BM455819	BM455819 AGENCOURT

5	891.6	37.6	1201	13 BX439033	BX439033 BX439033
6	876.2	37.0	943	13 BX342265	BX342265 BX342265
7	864.8	36.5	942	13 BX342264	BX342264 BX342264
8	858	36.2	867	13 BU521402	BU521402 AGENCOURT
9	842.6	35.6	936	13 BX678170	BX678170 AGENCOURT
10	839.8	35.4	924	13 BX330662	BX330662 BX330662
11	832	35.1	1201	13 BX345765	BX345765 BX345765
12	826.4	34.9	969	12 BX333256	BX333256 AGENCOURT
13	819.2	34.6	888	13 BX425968	BX425968 AGENCOURT
14	815.4	34.4	881	9 AL530006	AL530006 AL530006
15	812.2	34.3	884	9 AL530005	AL530005 AL530005
16	810.6	34.2	1070	12 BM460788	BM460788 AGENCOURT
17	804.6	33.9	942	13 BX677816	BX677816 AGENCOURT
18	803.6	33.9	859	13 BX680126	BX680126 AGENCOURT
19	803	33.9	1033	12 BX051239	BX051239 AGENCOURT
20	797	33.6	905	12 BG762465	BG762465 AGENCOURT
21	793.8	33.5	913	13 BX681982	BX681982 AGENCOURT
22	793.2	33.5	912	13 BX676527	BX676527 AGENCOURT
23	792.6	33.4	887	12 BX259163	BX259163 AGENCOURT
24	791	33.4	899	12 BG769533	BG769533 AGENCOURT
25	780.4	32.9	851	12 BX764996	BX764996 AGENCOURT
26	780.4	32.9	833	12 BX911625	BX911625 AGENCOURT
27	772.4	32.6	827	9 AU121914	AU121914 AU121914
28	772.4	32.6	949	13 BX691193	BX691193 AGENCOURT
29	769.4	32.5	1006	12 BM542149	BM542149 AGENCOURT
30	767.6	32.4	997	13 BX691213	BX691213 AGENCOURT
31	767.2	32.4	873	10 BG756781	BG756781 AGENCOURT
32	761	32.1	798	12 BX007147	BX007147 AGENCOURT
33	760.2	32.1	860	9 AU118064	AU118064 AU118064
34	757.8	32.0	798	12 BM973952	BM973952 AGENCOURT
35	756.6	31.9	909	12 BM006359	BM006359 AGENCOURT
36	752	31.7	773	13 BU008043	BU008043 AGENCOURT
37	749.6	31.6	765	12 BG763506	BG763506 AGENCOURT
38	748.6	31.6	792	11 BC013832	BC013832 Homo sapi
39	734.2	31.0	778	14 CD366834	CD366834 AGENCOURT
40	731.6	30.9	777	13 BX057631	BX057631 AGENCOURT
41	722.4	30.5	724	13 BM984976	BM984976 AGENCOURT
42	722.4	30.5	732	13 BX685777	BX685777 AGENCOURT
43	721.4	30.4	782	9 AU135766	AU135766 AGENCOURT
44	721.2	30.4	752	12 BM975862	BM975862 AGENCOURT
45	721.2	30.4	771	12 BM981191	BM981191 AGENCOURT

ALIGNMENTS

RESULT 1	AK086892	AK086892	2435 bp	mRNA	linear	HTC 05-DEC-2002
LOCUS	AK086892	AK086892	2435 bp	mRNA	linear	HTC 05-DEC-2002
DEFINITION	AK086892	AK086892	2435 bp	mRNA	linear	HTC 05-DEC-2002
ACCESSION	AK086892	AK086892	2435 bp	mRNA	linear	HTC 05-DEC-2002
VERSION	AK086892.1	AK086892.1	2435 bp	mRNA	linear	HTC 05-DEC-2002
KEYWORDS	HTC; CAP trapper.	HTC; CAP trapper.	2435 bp	mRNA	linear	HTC 05-DEC-2002
SOURCE	Mus musculus (house mouse)	Mus musculus (house mouse)	2435 bp	mRNA	linear	HTC 05-DEC-2002
ORGANISM	Mus musculus	Mus musculus	2435 bp	mRNA	linear	HTC 05-DEC-2002
REFERENCE	1	1	2435 bp	mRNA	linear	HTC 05-DEC-2002
AUTHORS	Carinci, P. and Hayashizaki, Y.	Carinci, P. and Hayashizaki, Y.	2435 bp	mRNA	linear	HTC 05-DEC-2002
TITLE	High-efficiency full-length cDNA cloning	High-efficiency full-length cDNA cloning	2435 bp	mRNA	linear	HTC 05-DEC-2002
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	Meth. Enzymol. 303, 19-44 (1999)	2435 bp	mRNA	linear	HTC 05-DEC-2002
MEDLINE	99279253	99279253	2435 bp	mRNA	linear	HTC 05-DEC-2002
PUBMED	10349636	10349636	2435 bp	mRNA	linear	HTC 05-DEC-2002
REFERENCE	2	2	2435 bp	mRNA	linear	HTC 05-DEC-2002
AUTHORS	Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	2435 bp	mRNA	linear	HTC 05-DEC-2002
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	2435 bp	mRNA	linear	HTC 05-DEC-2002
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	Genome Res. 10 (10), 1617-1630 (2000)	2435 bp	mRNA	linear	HTC 05-DEC-2002
MEDLINE	20499374	20499374	2435 bp	mRNA	linear	HTC 05-DEC-2002
PUBMED	11042159	11042159	2435 bp	mRNA	linear	HTC 05-DEC-2002


```

Db      933 GTTAGCC-TTTGGTGGAGCCGAGCTGACAGGACATCTTACAAGAAATTTTCACATCT 875
QY      1452 CTGGAAGCTTAGCAATCTATTGACACACTGTCGCTGGAA-TTTTGAAGACACATTC 1510
Db      874 CTGGAAGCTTAGCAATCTATTGACACACTGTCGCTGGAAATTTTGAAGACACATTC 815
QY      1511 TCCTCAGTGAAGCTAGAGCTTTTTCATTTTATTTTCTTCCATGCAAGCTGCTATCTCT 1570
Db      814 TCCTCAGTGAAGCTAGAGCTTTTTCATTTTATTTTCTTCCATGCAAGCTGCTATCTCT 755
QY      1571 GAAACGAGCTTAGAGTCCGCTTAGACGAGGAGGAGTTCGTAGAAAGCGGAC-C 1629
Db      754 GAAACGAGCGTATAGTCCGCTTAGACGAGGAGGAGTTCGTAGAAAGCGGAGCC 695
QY      1630 TGTCTTAAAGAGTCTCTCGACAGATCTGCTGGGCTGTGATGACGAATATTATGAATG 1689
Db      694 TGTCTTAAAGAGTCTCTCGACAGATCTGCTGGGCTGTGATGACGAATATTATGAATG 635
QY      1690 TGCCTTTTCTGAAGAGATTGTTGTAGCTCCAAAGCTTTTCTTATCCAGTGTTCAGTTC 1749
Db      634 TGCCTTTTCTGAAGAGATTGTTGTAGCTCCAAAGCTTTTCTTATCCAGTGTTCAGTTC 575
QY      1750 TTTATTTTCCCTTGTGATATGCTGTGTGAACCGTCTGTGAGTGTGATGCTGATCA 1809
Db      574 TTTATTTTCCCTTGTGATATGCTGTGTGAACCGTCTGTGAGTGTGATGCTGATCA 515
QY      1810 CAGATGGATTTTGTATTAAGCATCAATGTGACACTTGCAGGACACTACACGTGGACAT 1869
Db      514 CAGATGGATTTTGTATTAAGCATCAATGTGACACTTGCAGGACACTACACGTGGACAT 455
QY      1870 TGTGTGTTCTTCCATATTTTGAAGATAAATTTATGTGTAGACTTTTTTGTAAAGATACGG 1929
Db      454 TGTGTGTTCTTCCATATTTTGAAGATAAATTTATGTGTAGACTTTTTTGTAAAGATACGG 395
QY      1930 TTAATACATAAATTTATTAAGATGCTTCCATGACTCGTATTCAGATGCTTAAGAA 1989
Db      394 TTAATACATAAATTTATTAAGATGCTTCCATGACTCGTATTCAGATGCTTAAGAA 335
QY      1990 AGCATTTGCTGTACAAATATTTCTATTTTAGAAGGGTTTTATGACCAATGCCCCAG 2049
Db      334 AGCATTTGCTGTACAAATATTTCTATTTTAGAAGGGTTTTATGACCAATGCCCCAG 275
QY      2050 TTGTCAGTCAGAGCGTGTGTTTTCATTTTAAATGTCACCTGTAAATGGGCAT 2109
Db      274 TTGTCAGTCAGAGCGTGTGTTTTCATTTTAAATGTCACCTGTAAATGGGCAT 215
QY      2110 TATTTATGTTTTTTTTTGTGATTTCTGATTAATTTATGATTTATTAAGAAAGCTGTGT 2169
Db      214 TATTTATG-TTTTTTTTGTGATTTCTGATTAATTTATGATTTATTAAGAAAGCTGTGT 156
QY      2170 ACATTTGGGTTATACACTAGTATTTTAACTTACAGCTTATTTGTAATGTAACCCACC 2229
Db      155 ACATTTGGGTTATACACTAGTATTTTAACTTACAGCTTATTTGTAATGTAACCCACC 96
QY      2230 ATTTTAATGTAATTAATTAATGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 2289
Db      95 ATTTTAATGTAATTAATTAATGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 36
QY      2290 C-ACAACTTTTGTGCTGATANA 2314
Db      35 CNACAACTTTTGTGCTGATANA 10

```

```

RESULT 4
BM455819
LOCUS      BM455819
DEFINITION 1072 bp mRNA linear EST 05-FEB-2002
VERSION    AGENCOURT_5409165 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498261
ACCESSION  BM455819
VERSION     BM455819.1 GI:18504859
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

```

REFERENCE
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaaps-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The J.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the J.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM1219 row: 1 column: 06
            High quality sequence stop: 640.
FEATURES
source     1..1072
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5498261"
            /tissue_type="lymphoma, cell line"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_85"
            /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 1.867 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
BASE COUNT 282 a 277 c 250 g 262 t 1 others
ORIGIN

```

```

Query Match      38.7%; Score 916.4; DB 12; Length 1072;
Best Local Similarity 94.9%; Pred. No. 4.7e-101;
Matches 1011; Conservative 0; Mismatches 47; Indels 7; Gaps 6;

QY      8 GAGCGCTAACGCTCTTCTGTCTCCCGCGGTGTGATGAGCGTGAAGAACTGAGCGTCTA 67
Db      10 GAGCGCTAACGCTCTTCTGTCTCCCGCGGTGTGATGAGCGTGAAGAACTGAGCGTCTA 69
QY      68 AGGCGACCTCTTACTTACCTCAGGATGAGGCGTGTGGCAATTCATCGCTTCATGA 127
Db      70 AGGCGACCTCTTACTTACCTCAGGATGAGGCGTGTGGCAATTCATCGCTTCATGA 129
QY      128 AGCAGAGGAGGATGGGTCTGAACGACTTTATTGAGAGATTCGCAATTAACCTCTATGCAT 187
Db      130 AGCAGAGGAGGATGGGTCTGAACGACTTTATTGAGAGATTCGCAATTAACCTCTATGCAT 189
QY      188 GCAACACACCTTGAAGTTCAGTCCATCTTGAAGATCTCCCACTCAGAGCGCTGAGCTTA 247
Db      190 GCAACACACCTTGAAGTTCAGTCCATCTTGAAGATCTCCCACTCAGAGCGCTGAGCTTA 249
QY      248 TGAATGCCAACCCCTTCTCTCCACCAAGTCTCTTCAGCAAAATCAACCTTGGCCGCGT 307
Db      250 TGAATGCCAACCCCTTCTCTCCACCAAGTCTCTTCAGCAAAATCAACCTTGGCCGCGT 309
QY      308 CCAATCTTCATGCTTAACCATCTGACTTTCTACTTCTTGAAGATGATCGGAAGGCAAGTT 367
Db      310 CCAATCTTCATGCTTAACCATCTGACTTTCTACTTCTTGAAGATGATCGGAAGGCAAGTT 369
QY      368 TTGAAAGGTTCTTCTAGCAAGACACAGGCGAGAGAGTGTCTTATGCAAGTCAAAAGTTT 427
Db      370 TTGAAAGGTTCTTCTAGCAAGACACAGGCGAGAGAGTGTCTTATGCAAGTCAAAAGTTT 429
QY      428 TACAGAAGAAAGCAATCTCGAAAAAGAAAGAGAGAGCATATTATTCGAGCGGGAATG 487
Db      430 TACAGAAGAAAGCAATCTCGAAAAAGAAAGAGAGAGCATATTATTCGAGCGGGAATG 489
QY      488 TTCTGTTGAAGAAATGTGAAGCACCCTTTCTCTGTTGGGCTTCTACTTCTTTCAGACTG 547
Db      490 TTCTGTTGAAGAAATGTGAAGCACCCTTTCTCTGTTGGGCTTCTACTTCTTTCAGACTG 549

```

```

548 CTGACAAATCTACTTCTCTAGACATACATTAATGTTGGAGAGTTGTTCTACCATCTCC 607
550 CTGACAAATCTACTTCTCTAGACATACATTAATGTTGGAGAGTTGTTCTACCATCTCC 609
568 AGAGGAAAGCGTCTCTCTGAGACACAGGCTCTGTTTCTATGCTGCTGAATAGCCAGTG 667
570 AGAGGAAAGCGTCTCTCTGAGACACAGGCTCTGTTTCTATGCTGCTGAATAGCCAGTG 668
588 CTTTGGGCTACTCTGATCTCACTGAACATCTGTTATAGAGACTTAAACACAGACATATTT 727
590 CTTTGGGCTACTCTGATCTCACTGAACATCTGTTATAGAGACTTAAACACAGACATATTT 728
728 TGTAGATTCACAGGACACATCTGCTTACTGACTTCGAGCTCTGCAAGGAGAACCTT 786
729 TGTAGATTCACAGGACACATCTGCTTACTGACTTCGAGCTCTGCAAGGAGAACCTT 788
787 GAACACACAGACACATCTGCTTCTGT - GGCACGCCGAGTATCTCCACCTGAGGT 845
789 GAAC - CACACACACACATCTGCTTCTGTGGACCGCGAGTATCTCCACCTGAGGG 847
846 GCTTCATAGCAGCCTTATGACAGGACTGTGGAGCTGTGGCTGGGAGCTGTCTTGTGA 905
848 GCTTCATAGCAGCCTTATGACAGGACTGTGGAGCTGTGGCTGGGAGCTGTCTTGTGA 907
906 TGAGATGCTGTATGCTGCTGCTGCTGCTTCTTATAGCCGAAACACAGCTGAAATGTACGACAA 965
908 TGAAGAGCTGTATGCTGCTGCTGCTTCTTATAGCCGAAAC - ACCGCTGGAATGTCCGACAA 966
966 CATTCTGACACAGCCTCTGACGCTGTAACCAAAATATACAAATTCGCCACAGACACCTCT 1025
967 CTTCCGACACAGCCTCTGACGCTGTAACCAAAATATACAAATTCGCCACAGACACCTCT 1026
1026 CGAGGCGCTCC--TGCAGAGGACAGGACAAAGCGGCTCGGGGCC 1068
1027 GGAAGGCCCTCTCGCAAGAGGAAAGGAAAGGCGCTCGGGGCC 1071

RESULT 5
BX439033 1201 bp mRNA linear EST 15-MAY-2003
LOCUS
DEFINITION
BX439033 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YJ09
5-PRIME, mRNA sequence.
ACCESSION
BX439033
VERSION
BX439033.1 GI:30783742
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequençage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3341.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE005CE05Qp1&cluster=3341.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005CE05Qp1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YJ09"
/tissue_type="PLACENTA"
FEATURES
source

```

```

/clone.lib="Homo sapiens PLACENTA"
[Note]-Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized.
BASE COUNT 316 a 271 c 279 g 296 t 39 others
ORIGIN
Query Match 37.6%; Score 891.6; DB 13; Length 1201;
Best Local Similarity 98.7%; Pred. No. 4.1e-98;
Matches 916; Conservative 3; Mismatches 7; Indels 2; Gaps 2;
QY 8 GAGCGCTAACGCTCTTCTCTCCCGCGGTGGTGGATGAGCGGTGAAACAGGAGCGTCTA 67
DB 71 GAGCGCTAACGCTCTTCTCTCCCGCGGTGGTGGATGAGCGGTGAAACAGGAGCGTCTA 130
QY 68 AGGCGACCCCTCACCTTACTCCAGGATGAGGGGCGATGGTGGCAATTTCTCATGCTTTCATGA 127
DB 131 AGGCGACCCCTCACCTTACTCCAGGATGAGGGGCGATGGTGGCAATTTCTCATGCTTTCATGA 190
QY 128 AGCAGAGGAGGATGGTCTGAAACGACTTTATTACAGAAGATTGCCAATAATCTCTATGCAAT 187
DB 191 AGCAGAGGAGGATGGTCTGAAACGACTTTATTACAGAAGATTGCCAATAATCTCTATGCAAT 250
QY 188 GCAAAACACCCCTGAACTTTCAGTCCCATCTTGAAGATCTCCCAACCTTCAGAGGCGCTGA 247
DB 251 GCAAAACACCCCTGAACTTTCAGTCCCATCTTGAAGATCTCCCAACCTTCAGAGGCGCTGA 310
QY 248 TGAATGCCAACCCCTTCTCTCCAGCAAGTCTCTCTCAGCAAAATCAACCTTGGCGCTCTGCT 307
DB 311 TGAATGCCAACCCCTTCTCTCCAGCAAGTCTCTCTCAGCAAAATCAACCTTGGCGCTCTGCT 370
QY 308 CCAATCCCTCATGCTTAAACCATCTGACTTTTCTTGAAGATGATCGGAAAGGCGAGTT 367
DB 371 CCAATCCCTCATGCTTAAACCATCTGACTTTTCTTGAAGATGATCGGAAAGGCGAGTT 430
QY 368 TTGGAAAGGTTCTTCTAGCAGACACAAAGGCGAGAGTGTCTTATGCAATCAAGATTT 427
DB 431 TTGGAAAGGTTCTTCTAGCAGACACAAAGGCGAGAGTGTCTTATGCAATCAAGATTT 490
QY 428 TACAGAAGAAGCAATCTCTGAAAGAAAGAGAGAGCATATTATGTCGGAGCGCAATG 487
DB 491 TACAGAAGAAGCAATCTCTGAAAGAAAGAGAGAGCATATTATGTCGGAGCGCAATG 550
QY 488 TTCTGTTGAAGATGTAAGCAACCGTTTCTTGGTGGGCTTCACTTCTCTTCCAGACATG 547
DB 551 TTCTGTTGAAGATGTAAGCAACCGTTTCTTGGTGGGCTTCACTTCTCTTCCAGACATG 610
QY 548 CTGCAAAATTTGACTTTTCTCTAGACTACATTAATGTTGGAGAGTGTCTTACCATCTCC 607
DB 611 CTGCAAAATTTGACTTTTCTCTAGACTACATTAATGTTGGAGAGTGTCTTACCATCTCC 670
QY 608 AGAGGAAAGCGTCTCTCTGGAACACCGGCTCTCTTCTATGCTGCTGAATAGCCAGTG 667
DB 671 AGAGGAAAGCGTCTCTCTGGAACACCGGCTCTCTTCTATGCTGCTGAATAGCCAGTG 730
QY 668 CTTTGGGCTACTCTGATCTCACTGAACATCTGTTTATAGAGACTTAAACACAGACATATTT 727
DB 731 CTTTGGGCTACTCTGATCTCACTGAACATCTGTTTATAGAGACTTAAACACAGACATATTT 790
QY 728 TGTAGATTCACAGGACACATCTGCTTACTGATTTTGGAGCTCTGCAAGGAGAACATG 787
DB 791 TGTAGATTCACAGGACACATCTGCTTACTGACTTCGAGCTCTGCAAGGAGAACATG 850
QY 788 AACCAACACAGACACATCTGCTTCTGTCGCGCGGAGTATCTCGACCTCGAGGTGC 847
DB 851 AACCAACAGACACATCTGCTTCTGTCGCGCGGAGTATCTCGACCTCGAGGTGC 910
QY 848 TTCATAAGCAGCCTTATGACAGGACTGTGGCTGGTGGGAGCTGTCTTGTATG 907
DB 911 TTCATAAGCAGCCTTATGACAGGACTGTGGCTGGTGGGAGCTGTCTTGTATG 969

```



```

QY 908 AGATGCTGTATGCTGCGCGCCCTTTTATA 935
Db 970 ARAT-SIGTATGGGCTGGCGCTTTTATA 996

RESULT 6
BX342265 943 bp mRNA linear EST 01-MAY-2003
LOCUS BX342265 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK011YB02 5-PRIME, mRNA sequence.
ACCESSION BX342265
VERSION BX342265.1 GI:30311068
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3341.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK011DA01Q1P1
Feng Liang Email: filiang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0DK011DA01Q1P1.
Location/Qualifiers
1..943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YB02"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 254 a 226 c 218 g 241 t 4 others
ORIGIN
Query Match 37.0%; Score 876.2; DB 13; Length 943;
Best Local Similarity 99.3%; Pred. No. 3.4e-96;
Matches 875; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 GGGAGCGCTAACGCTTCTGCTCCCGCGGGTGATGACGGTGAAGAGCTGCTGC 65
Db 63 GGGATCGCTAACGCTTCTGCTCCCGCGGGTGATGACGGTGAAGAGCTGCTGC 122

QY 66 TAAGGACCGCTCACTTACTCCAGATGAGGCGATGCTGGCAATTCTATCGCTTTCAT 125
Db 123 TAAGGACCGCTCACTTACTCCAGATGAGGCGATGCTGGCAATTCTATCGCTTTCAT 182

QY 126 GAAGCAGAGGAGGATGGTGTGACGACTTTATTCAGAGATTGCCAATAACTCCATGCG 185
Db 183 GAAGCAGAGGAGGATGGTGTGACGACTTTATTCAGAGATTGCCAATAACTCCATGCG 242

QY 186 ATGCAACACCGCTCAAGTTAGTCCATCTGAGAGTCTCCCAACCTCAGAGCTGAGCT 245
Db 243 ATGCAACACCGCTCAAGTTAGTCCATCTGAGAGTCTCCCAACCTCAGAGCTGAGCT 302

QY 246 TATGAATGCCAACCGCTTCTCTCCACCAAGTCTTCTCAGCAATAAATCGGCGCGTC 305
Db 303 TATGAATGCCAACCGCTTCTCTCCACCAAGTCTTCTCAGCAATAAATCGGCGCGTC 362

FEATURES
SOURCE
1..943
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YB02"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 254 a 226 c 218 g 241 t 4 others
ORIGIN
Query Match 37.0%; Score 876.2; DB 13; Length 943;
Best Local Similarity 99.3%; Pred. No. 3.4e-96;
Matches 875; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 6 GGGAGCGCTAACGCTTCTGCTCCCGCGGGTGATGACGGTGAAGAGCTGCTGC 65
Db 63 GGGATCGCTAACGCTTCTGCTCCCGCGGGTGATGACGGTGAAGAGCTGCTGC 122

QY 66 TAAGGACCGCTCACTTACTCCAGATGAGGCGATGCTGGCAATTCTATCGCTTTCAT 125
Db 123 TAAGGACCGCTCACTTACTCCAGATGAGGCGATGCTGGCAATTCTATCGCTTTCAT 182

QY 126 GAAGCAGAGGAGGATGGTGTGACGACTTTATTCAGAGATTGCCAATAACTCCATGCG 185
Db 183 GAAGCAGAGGAGGATGGTGTGACGACTTTATTCAGAGATTGCCAATAACTCCATGCG 242

QY 186 ATGCAACACCGCTCAAGTTAGTCCATCTGAGAGTCTCCCAACCTCAGAGCTGAGCT 245
Db 243 ATGCAACACCGCTCAAGTTAGTCCATCTGAGAGTCTCCCAACCTCAGAGCTGAGCT 302

QY 246 TATGAATGCCAACCGCTTCTCTCCACCAAGTCTTCTCAGCAATAAATCGGCGCGTC 305
Db 303 TATGAATGCCAACCGCTTCTCTCCACCAAGTCTTCTCAGCAATAAATCGGCGCGTC 362

RESULT 7
BX342264 942 bp mRNA linear EST 01-MAY-2003
LOCUS BX342264 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DK011YB02 3-PRIME, mRNA sequence.
ACCESSION BX342264
VERSION BX342264.1 GI:30308908
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3341.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DK011DA01NP1
Feng Liang Email: filiang@lifetech.com
URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CS0DK011DA01NP1.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/clone="CS0DK011yB02"
/cell_line="HELA CELLS COT 25-NORMALIZED"
/cell_type="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      375 a 193 c 171 g 238 t 15 others
ORIGIN

Query Match      36.58; Score 864.8; DB 13; Length 942;
Best Local Similarity 99.94; Pred. No. 7.9e-95;
Matches 907; Conservative 13; Mismatches 12; Indels 4; Gaps 4;

QY 1397 CTTTTTGGTGAGCGCCGACGCTGACAGGACATCTTACAAGAAATTTGCACATCTCTGA 1456
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 942 SCTTTTGGTGAGCGCCGACGCTGACAGGACATCTTACAAGAAATTTGCACATCTCTGA 883
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1457 AGCTTAGCAATCTTATTGGACACGTTCGCTGAA-TTTTGAAGACACATCTCTCC 1515
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 882 AGCTTAGCAATCTTATTGGACACGTTCGCTGAA-TTTTGAAGACACATCTCTCC 823
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1516 AGTGAGCTCARGAGGTTTTCATTTTATCTCTCTCCAAAGTGGTCTATCTGAAC 1575
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 822 AGTGAGCTCARGAGGTTTTCATTTTATCTCTCTCCAAAGTGGTCTATCTGAAC 763
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1576 GAGCTTAGAGTGGCGCTTAGACGAGGAGGAGTTTCGTTAGAAGCGGAC-CTGTTC 1634
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 762 GAGCTTAGAGTGGCGCTTAGACGAGGAGGAGTTTCGTTAGAAGCGGACCGTTC 703
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1635 TAAAGAGGTCCTGCGAGTCTCTGCGGCTGTGATGACGAATATATGAATGCGCT 1694
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 702 TAAAGAGGTCCTGCGAGTCTCTGCGGCTGTGATGACGAATATATGAATGCGCT 643
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1695 TTCTGAGAGATGTTGTAGCTCCAAAGCTTTTCCTATCGAGTGTTCAGTCTTTAT 1754
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 642 TTCTGAGAGATGTTGTAGCTCCAAAGCTTTTCCTATCGAGTGTTCAGTCTTTAT 583
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1755 TTCTCCTTGTGGATATGCTGTGAACGCTGCTGTGAGTGTGTATGCCGTATCAGAT 1814
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 582 TTCTCCTTGTGGATATGCTGTGAACGCTGCTGTGAGTGTGTATGCCGTATCAGAT 523
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1815 GGATTTTGTATAAGCAATCAATGTGACACTTGCAGGACACATAGCAAGTGGGACATTTGTT 1874
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 522 GGATTTTGTATAAGCAATCAATGTGACACTTGCAGGACACATAGCAAGTGGGACATTTGTT 463
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1875 GTTCTTCCATATTTGGAGATAATTTATGCTGTAGACTTTTGTGAAGATACGGTAA 1934
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 462 GTTCTTCCATATTTGGAGATAATTTATGCTGTAGACTTTTGTGAAGATACGGTAA 403
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1935 AACTAAATTTATTGAATGCTTGTGAATGACCTGATTCAGATGCGCTAAAGAAAGCAT 1994
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 402 AACTAAATTTATTGAATGCTTGTGAATGACCTGATTCAGATGCGCTAAAGAAAGCAT 343
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1995 TGCTGCTACAATATTTCTATTTTGAAGAGGTTTATGGACCAATGCCAGTGTG 2054
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 342 TGCTGCTACAATATTTCTATTTTGAAGAGGTTTATGGACCAATGCCAGTGTG 283
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 2055 AGTCAGAGCGGTTGGTGTGTTTCAATGTTTAAATGTCACTGCTAAATGGGCATTTT 2114
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 282 AGTCAGAGCGGTTGGTGTGTTTCAATGTTTAAATGTCACTGCTAAATGGGCATTTT 223
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 2115 ATGTTTTTTTTCATCTCTGTAATTTGATGATGATGATGATGATGATGATGATGAT 2174
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 222 ATG-TTTTTTTTTCATCTCTGTAATTTGATGATGATGATGATGATGATGATGATGAT 164
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 2175 GGTTTATAACACTAGTATATTTAAACTTACAGGCTTATTTGTAATGTAACACCACTTTT 2234
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 GGTTTATAACACTAGTATATTTAAACTTACAGGCTTATTTGTAATGTAACACCACTTTT 104
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 2235 ATGTTACTGTATTTACATGGTTTAAATACGTACAACTTCCCTCATCCCATCACAA 2294
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

103 AATGTACTGTAAATTAACATGGTTATATACGTACATCTTCCCTCATCCATCACAA 44
2295 CTTTTTTTGTGTGTGATAAACTGATTTTGGTTTGA 2330
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43 CTTTTTTTGTGT-TGATAAACTGATTTTGGTTNNCH 9

RESULT 8
BU521402
LOCUS      BU521402
DEFINITION AGENCOURT_10167704 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6527249
5' mRNA sequence.
ACCSSION   BU521402
VERSION    BU521402.1 GI:22828928
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 867)
AUTHORS     NIH-MGC http://mgc.ncl.nih.gov/
JOURNAL     Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLNL4123 row: n column: 17
            High quality sequence stop: 735.
FEATURES
     source
     1..867
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6527249"
     /tissue_type="leiomyosarcoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH_MGC_71"
     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
     Average insert size 2.1 kb."
BASE COUNT 233 a 212 c 196 g 226 t
ORIGIN

Query Match      36.28; Score 858; DB 13; Length 867;
Best Local Similarity 99.4%; Pred. No. 5.5e-94;
Matches 861; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 13 CTAACGCTTTTCTCTCCCGCGGTGGTGTGATGACGGTGAAAACCTGAGGCTCTAAGGCG 72
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 CTAACGCTTTTCTCTCTCCCGCGGTGGTGTGATGACGGTGAAAACCTGAGGCTCTAAGGCG 60
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 73 ACCCTCACTTACTCCAGGATGAGGGGCAATGCTCATCGCTTTCATGAGGAG 132
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ACCCTCACTTACTCCAGGATGAGGGGCAATGCTCATCGCTTTCATGAGGAG 120
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 133 AGGAGGATGGCTCGAACGACTTTATTGAGAACATTGCCAATAACCTTCATCATGAGAA 192
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 AGGAGGATGGCTCGAACGACTTTATTGAGAACATTGCCAATAACCTTCATCATGAGAA 180
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 193 CACCTCAAGTTCAGTCCATCTTTGAAGATCTCCACACCTCAGAGGCTTGAGCTTATGAT 252
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 CACCTCAAGTTCAGTCCATCTTTGAAGATCTCCACACCTCAGAGGCTTGAGCTTATGAT 240
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 253 GCGAACCTTCTCTCCACCAAGCTCTTCACCAATCAACCTTGGCCGCTGCTCCAT 312
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 GCGAACCTTCTCTCTCCACCAAGCTCTTCACCAATCAACCTTGGCCGCTGCTCCAT 300
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```


Db 781 CATTGCTGCTANCAATATTTCTATTTTATAGAAAGGTTTATGGAACCAATGCCCGCT 840
 QY 2051 TGTGAGCAG-AGCGGTGGTGGTTTCATGTTTAAATGTCACCTGTAAATGGCA- 2108
 Db 841 TGTGAGTCAAAACCGTGGGGTTTTCATGTTTAAATGTCACCTGTAAATGGGAC 900
 QY 2109 TTAATTATGTTTTTTTTTTCGATCTCTGATAATG 2144
 Db 901 TAATTAGGCTTTTTTTTTTGCCCTCTGGAATTTG 936

RESULT 10

BX330662/c
 LOCUS
 DEFINITION
 clone CS0D1078Y023 3-PRIME, mRNA sequence.

ACCESSION

BX330662

VERSION

BX330662.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens (human)

REFERENCE

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3341.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0BAK064DC12NM1&cluster=3341.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ InvitroSen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0BAK064DC12NM1.

Location/Qualifiers

1. 924

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0B1078Y023"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 337 a 179 c 162 g 246 t

ORIGIN

Query Match 35.4%; Score 839.8; DB 13; Length 924;

Best Local Similarity 97.7%; Pred. No. 8e-92;

Matches 894; Conservative 0; Mismatches 17; Indels 4; Gaps 4;

QY 1435 AGAGAATTGACATCTCTGGAAGCTTAGCAATCTTAATGCA-CACGTGTCGTGGAA-T 1492

Db 917 AGAAATTTGGCCCTTCTTGAAGCTTAGCAATCTTATTCACCACTGTCGCGTGAAGC 858

QY 1493 TTTTGAAGACACATCTCTCAGTCAGCTCATGAGGTTTTCATTTTATCTT-CCTT 1551

Db 857 TTTTGAAGACACATCTCTCAGTCAGCTCATGAGGTTTTCATTTTATCTTCCCTT 798

QY 1552 CCAAGCTGCTATCTCTGAAACAGGCTTAGAGTCCGCTTAGAGCGGACGAGT 1611

Db 797 CCAAGCTGCTATCTCTGAAACAGGCTTAGAGTCCGCTTAGAGCGGACGAGT 738

QY 1612 TTCGTAGAAACGGAC-CTGTTCTAAAAGGCTCTCTCAGATCTGCTGGCTGTGA 1670

Db 737 TTCGTAGAAACGGACGCTGTCTTAAAAAAGGCTCTCTCAGATCTGCTGGCTGTGA 678

QY 1671 TGAGCAATATTATGAATGTCCTTTCTGAGAGATTTGTTAGTCTCAAGCTTTTCC 1730
 Db 677 TGACCAATATTATGAATGTCCTTTCTGAGAGATTTGTTAGTCTCAAGCTTTTCC 618
 QY 1731 TATGCGAGTGTTCAGTTCCTTTATTTCCCTGCTGATATGCTGTGTAACCGTCTGTG 1790
 Db 617 TATGCGAGTGTTCAGTTCCTTTATTTCCCTGCTGATATGCTGTGTAACCGTCTGTG 558
 QY 1791 AGTGTGTATGCTGATCAGATGATTTGTTTATAAGCATCAATGTGACATTTGCGAG 1850
 Db 557 AGTGTGTATGCTGATCAGATGATTTGTTTATAAGCATCAATGTGACATTTGCGAG 498
 QY 1851 ACATACAAGCTGGGACATTTGTTTCTTCCATATTTGGAAGATAAATTTATGTGTAG 1910
 Db 497 ACATACAAGCTGGGACATTTGTTTCTTCCATATTTGGAAGATAAATTTATGTGTAG 438
 QY 1911 ACTTTTTTGAAGATACGGTTTAATACTAAATTTTGAATGTTGTCATACACATCG 1970
 Db 437 ACTTTTTTGAAGATACGGTTTAATACTAAATTTTGAATGTTGTCATACACATCG 378
 QY 1971 TATTCAGATGCTTAAGAAAGCATTTGCTGCTACAAATTTCTATTTTAGAAAGGTTT 2030
 Db 377 TATTCAGATGCTTAAGAAAGCATTTGCTGCTACAAATTTCTATTTTAGAAAGGTTT 318
 QY 2031 TTATGGACCAATGCCAGTGTGTCAGTCAGAGCGTGTGTTTTCATTTTAAAGT 2090
 Db 317 TTATGGACCAATGCCAGTGTGTCAGTCAGAGCGTGTGTTTTCATTTTAAAGT 258
 QY 2091 TCACCTGTAAATGGGCATTTATTTGTTTTTTTTTTCGATTCCTGATAATGTATGTA 2150
 Db 257 TCACCTGTAAATGGGCATTTATTTGTTTTTTTTTTCGATTCCTGATAATGTATGTA 198
 QY 2151 TTGTATAAAGAACCTCTGTACATTTGGTTTATAACACTAGTATATTTAAACTTACAGGCTT 2210
 Db 197 TTGTATAAAGAACCTCTGTACATTTGGTTTATAACACTAGTATATTTAAACTTACAGGCTT 138
 QY 2211 ATTTGTAATGTAAGACACCTTTTAATGTAAGTCTGTAATTAACAGTGTATTAACAGTACAA 2270
 Db 137 ATTTGTAATGTAAGACACCTTTTAATGTAAGTCTGTAATTAACAGTGTATTAACAGTACAA 78
 QY 2271 TCCTTCCCTCATCCCATCACAACTTTTTTTTGTGTGTAATAACTGATTTTGGTTTGA 2330
 Db 77 TCCTTCCCTCATCCCATCACAACTTTTTTTGTGTGTAATAACTGATTTTGGTTTGA 18
 QY 2331 ATAAACCTTGAAA 2345
 Db 17 ATAAACCTTGAAA 3

RESULT 11

BX345765

LOCUS

BX345765

DEFINITION

clone CS0D104YU21 5-PRIME, mRNA sequence.

ACCESSION

BX345765

VERSION

BX345765.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li.W.B., Gruber.C., Jesse,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3341.r For

more information about this cluster, see

http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0A1040Cf1l10p1&cluster=3341.r. Contact :
Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0A1040Cf1l10p1.
Location/Qualifiers

FEATURES
source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0A1040Y21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 301 a 277 c 202 g 379 t 42 others
ORIGIN

Query Match 35.1a; Score 832; DB 13; Length 1201;
Best Local Similarity 92.8a; Pred. No. 5.7e-91;
Matches 903; Conservative 10; Mismatches 56; Indels 4; Gaps 4;
QY 9 AGCGCTAACCTCTTCTGCTCCCGCGGTGGTGTGATGAGGTCGAAACCTGAGGCTGCTAA 68
DB 77 ATCTCTACATCTTCTTCTCCGCTCGGTGGTATTACGGTTAAACCTTANGCTCTAA 136
QY 69 GGGCACCTCCTTACTTCCAGGATGAGGGGCATGTTGGCAATCTCTATCGCTTTCATGAA 128
DB 137 GGGCACCTCCTTACTTCCAGGATGAGGGGCATGTTGGCAATCTCTATCGCTTTCATGAA 196
QY 129 GCAGGAGGATGGCTTGAACGACTTATTAGAGATTCGCAATCACTCTATGCTGATG 188
DB 197 GCAGGAGGATGGCTTGAACGACTTATTAGAGATTCGCAATCACTCTATGCTGATG 256
QY 189 CAACACCTTGAATCTCAGTCCATCTTGAAGATCTCCCACTCAGGAGCTGAGCTTAT 248
DB 257 CTAAACCTTAAATTTTCACTCTTCACTCTTCACTCTTCACTCTTCACTCTTCACTCT 316
QY 249 GAATGCCAACCTTCTCTCCACCAAGTCTCTCAGCAATCACTCTGCGCGCTGCTG 308
DB 317 TAATCCACCTTCTCTCCACCAAGTCTCTCAGCAATCACTCTGCGCGCTGCTG 376
QY 309 CAATCCTCATCTTAACATCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 368
DB 377 CAATCCTCATCTTAACATCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 436
QY 369 TGGAAAGCTTCTCTAGCAACACAAAGGAGAGAGAGTCTCTATGCACTCAAGTCTT 428
DB 437 TGAAGTCTCTCTAGCAACACAAAGGAGAGAGTCTCTATTCAGTCAATATTTT 496
QY 429 ACAGAAAGACATCTCTGAAAGAGAGAGAGAGATTAATGTCGAGGAGGAGATGT 488
DB 497 ACAGAAAGATCTCTCTGAAAGAGAGAGAGAGATTAATGTCGAGGAGGAGATGT 556
QY 489 TCTGTGAAGATGTGAAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 548
DB 557 TCTTTGAAGATGTGAAGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 616
QY 549 TGACAAATTTGACTTGTCT 608
DB 617 TTACAAATTTGACTTGTCT 676
QY 609 GAGGAGAGCTGCT 668
DB 677 GAGGAGAGCTGCT 736
QY 669 CTGCGCTACCTGCT 728
DB 737 CTGCGCTACCTGCT 796
QY 729 GCTAGATTCACAGGAGACATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 788

DB 797 GCTAGATTCACAGGAGACATTTGCTTACTCAGTTCGAGCTCTCAGAGGAGACATGGA 856
QY 789 ACACACAGCAGACAAATCCACCTTCTGTCGACGCGGAGAGTATCTCGCACCTGA-GSTGC 847
DB 857 ACACACAGCAGACAAATCCACCTTCTGTCGACGCGGAGAGTATCTCGCACCTGAGGTTST 916
QY 848 TTCTATAGCAGCCTTATGACAGG-ACTGTGAGTGTGTGCTGCTGGGAGCTCTCTTGTAT 906
DB 917 TTCTATAGCAGCCTTATGACAGGAGTGTGAGTGTGTGCTGCTGGGAGCTCTCTTGTAT 976
QY 907 GAGATCTGTATGCTGCTGCGCTTTTATAGCCGAAACACAGCTGAAATGTAGCACAC 966
DB 977 GAGATCTGTATGCTGCTGCGCTTTTATAGCCG-AAACACAGCTGAAATGTAGCAC-AC 1034
QY 967 ATTCTGAACAGC 979
DB 1035 ATTCTGAACAC 1047

RESULT 12
BI333256
LOCUS
DEFINITION 60296596f1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5138951 5',
mRNA sequence.
ACCESSION BI333256
VERSION BI333256.1 GI:15017913
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 969)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11341 row: p column: 24
High quality sequence stop: 885.
Location/Qualifiers
1. 969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5138951"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1; NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

BASE COUNT 236 a 255 c 229 g 249 t
ORIGIN
Query Match 34.9a; Score 826.4; DB 12; Length 969;
Best Local Similarity 98.2a; Pred. No. 3.1e-90;
Matches 878; Conservative 0; Mismatches 11; Indels 5; Gaps 4;
QY 632 CACGGCTCGTTTCTATGCTCTCTGAAATAGCCAGTGCCTTGGGCTACCTGCTACTGTA 691
DB 1 CACGGCTCGTTTCTATGCTCTCTGAAATAGCCAGTGCCTTGGGCTACCTGCTACTGTA 60
QY 692 ACATCGTTTATAGAGACTTAAACACAGAGATATTTTGTAGATTTCACAGGAGACATG 751

Db 61 ACATCGTTTATAGAGACTTAAACCACAGAGATATTTTGTAGATTCACAGGACACATTTG 120

Qy 752 TCCTTACTGATTTTGGAGCTCTGCAAGGAGAACATTGAACACACAGCAGCAACATCCACT 811

Db 121 TCCTTACTGATTTTGGAGCTCTGCAAGGAGAACATTGAACACACAGCAGCAACATCCACT 180

Qy 812 TCTGTGGCAGCCGGAGTATCTGCGACCTGAGTGTCTTATAGAGAGCCCTTATGACAGGA 871

Db 181 TCTGTGGCAGCCGGAGTATCTGCGACCTGAGTGTCTTATAGAGAGCCCTTATGACAGGA 240

Qy 872 CTGTGGAGCTGTGTGCTGGAGCTGTCTTGTATGAGATGCTGTATGGCCCTGCGCCCTT 931

Db 241 CTGTGGAGCTGTGTGCTGGAGCTGTCTTGTATGAGATGCTGTATGGCCCTGCGCCCTT 300

Qy 932 TTTATAGCCGAACACAGCGGAAATATAGACACACATCTGAACAGAGCCCTTCCAGCTGA 991

Db 301 TTTATAGCCGAACACAGCGGAAATATAGACACACATCTGAACAGAGCCCTTCCAGCTGA 360

Qy 992 AACCAAAATATTACAAATTCGCAAGACACTCTCTGGAGGCCCTCTGCAAGAGGACAGGA 1051

Db 361 AACCAAAATATTACAAATTCGCAAGACACTCTCTGGAGGCCCTCTGCAAGAGGACAGGA 420

Qy 1052 CAAAGCGCTCGGGGCCAAGGATGACTTTCATGAGAGATTAGAGTCAATGCTTCTTCTCT 1111

Db 421 CAAAGCGCTCGGGGCCAAGGATGACTTTCATGAGAGATTAGAGTCAATGCTTCTTCTCT 480

Qy 1112 TAATTAAGTGGGATGATCTCATTAAAGAGATTAATCTCCCTCTTTTAAACCAAAATGTA 1171

Db 481 TAATTAAGTGGGATGATCTCATTAAAGAGATTAATCTCCCTCTTTTAAACCAAAATGTA 540

Qy 1172 GTGGGCCCAACGAGCTACGACATTTGACCCGAGTTTACCAGAGAGCCCTGTCGCCAAT 1231

Db 541 GTGGGCCCAACGAGCTACGACATTTGACCCGAGTTTACCAGAGAGCCCTGTCGCCAAT 600

Qy 1232 CCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGAGAGCTGCGAGG 1291

Db 601 CCATTGCAAGTCCCTGACAGCGTCTCTGTCACAGCCAGCGTCAAGAGAGCTGCGAGG 560

Qy 1292 CTTTCCAGAGCTTTTCTATGCGGCTGCGACAGGACTTCTTCTGCAAGCCCTGTTAGGCG 1351

Db 661 CTTTCCAGAGCTTTTCTATGCGGCTGCGACAGGACTTCTTCTGCAAGCCCTGTTAGGCG 720

Qy 1352 TTGGTTTAA--GGATTATATGTTTATGTTTCCCAATGTTTATGTTAGCCCTTTTGGTGGAG 1409

Db 721 TTGGTTTAA--GGATTATATGTTTATGTTTCCCAATGTTTATGTTAGCCCTTTTGGTGGAG 780

Qy 1410 CCGCCAGCTGACAGGAGATCTTACAGAGAAATTTGACACA-TCTCTGGAAGCTTAGCAATC 1468

Db 781 CCG-CAGCTGACAGGAGATCTTACAGAGAAATTTGACACA-TCTCTGGAAGCTTAGCAATC 839

Qy 1469 TTATGACACT-GTTGCTGGAATTTTGAAGAGCAGCATCTCTCTCAGTGAG 1521

Db 840 TTATGACACTGTTGCTGGAATTTTGAAGAGCAGCATCTCTCTCAGTGAG 893

RESULT 13

LOCUS BQ425968

DEFINITION AGENCOURT_7919608 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6160295

ACCESSION BQ425968

VERSION BQ425968.1 GI:21121283

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 888)

TITLE NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/OTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
Plate: LIML3510 row: 1 column: 24
High quality sequence stop: 699.

FEATURES

Source

1..888

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6160295"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1; Nucleotide: 2; Salt: Cloned unidirectionally. Primer: oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 235 a 218 c 207 g 227 t 1 others

ORIGIN

Query Match 34.6%; Score 819.2; DB 13; Length 888;

Best Local Similarity 99.6%; Pred. No. 2.4e-89;

Matches 821; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 54 AACTGAGGCTGCTAAGGGCACCCCTCACTTACTCCAGATGAGGGGATGCTGCAATTC 113

Db 1 AACTGAGGCTGCTAAGGGCACCCCTCACTTACTCCAGATGAGGGGATGCTGCAATTC 60

Qy 114 CATCGCTTTCATGAAGCAGAGGAGGATGGTCTGAACGACTTATTTCAGAAATTC 173

Db 61 CATCGCTTTCATGAAGCAGAGGAGGATGGTCTGAACGACTTATTTCAGAAATTC 120

Qy 174 TAATCTCTATGATGCAAGAACACCCCTGAAGTTCAGTCCATCTTGAAGATCTCTCAACCTCA 233

Db 121 TAATCTCTATGATGCAAGAACACCCCTGAAGTTCAGTCCATCTTGAAGATCTCTCAACCTCA 180

Qy 234 GGAGCCTGAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTTCTCAGCAAAATCA 293

Db 181 GGAGCCTGAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTTCTCAGCAAAATCA 240

Qy 294 CTTTGGCCGCTGCTCAATCTCTATGCTAAACCATCTGACTTTTCACTTCTTGAAGTAT 353

Db 241 CTTTGGCCGCTGCTCAATCTCTATGCTAAACCATCTGACTTTTCACTTCTTGAAGTAT 300

Qy 354 CGGAAAGGGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAAAGGAGAGAGAGATGTTCTA 413

Db 301 CGGAAAGGGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAAAGGAGAGAGATGTTCTA 360

Qy 414 TGCAGTCAAGTCTTTCAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGAGATATAT 473

Db 361 TGCAGTCAAGTCTTTCAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGATATAT 420

Qy 474 GTCCGAGCGGAATCTCTGTTGAAGATGTTGAAGCAACCCCTTCTCTGCGGCTTCACT 533

Db 421 GTCCGAGCGGAATCTCTGTTGAAGATGTTGAAGCAACCCCTTCTCTGCGGCTTCACT 480

Qy 534 CTCTTCCAGACTGCTGACAAATTTGACTTTTGTCTAGACTACATTAATGTTGGAGATT 593

Db 481 CTCTTCCAGACTGCTGACAAATTTGACTTTTGTCTAGACTACATTAATGTTGGAGATT 540

Qy 594 GTTCTACCATCTCCAGAGGAGAGCGTCTTCTGGAACACAGCGGCTGCTTCTATGCTGC 653

Db 541 GTTCTACCATCTCCAGAGGAGAGCGTCTTCTGGAACACAGCGGCTGCTTCTATGCTGC 600

Qy 654 TGAATAGCAGCTGCTTGGCTTACCTGCAATTCACATGCTTATAGAGACTTAA 713

Db 601 TGAATAGCAGCTGCTTGGCTTACCTGCAATTCACATGCTTATAGAGACTTAA 660

```

QY 714 ACCAGAGATATTTCTAGATTACAGGAGACACATGTCCTTACTGATTTCGGACTCTG 773
DB 661 ACCAGAGATATTTCTAGATTACAGGAGACACATGTCCTTACTGACTTCGGACTCTG 720
QY 774 CAAGGAGACATTTACACACAGCAGACACATCCACCTCTCTGTCGCGGAGTACT 833
DB 721 CAAGGAGACATTTACACACAGCAGACACATCCACCTCTCTGTCGCGGAGTACT 780
QY 834 CGCACCTGAGCTCTTCAAGCAGCCTTATGACAGGACTGTGG 877
DB 781 CGCACCTGAGCTCTTCAAGCAGCCTTATGACAGGACTGTGG 824

RESULT 14
LOCUS AL530006 881 bp mRNA linear EST 23-MAY-2003
DEFINITION AL530006 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD009YH16 5-PRIME, mRNA sequence.
ACCESSION AL530006
VERSION AL530006.2 GI:31067844
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
On Feb 13, 2001 this sequence version replaced gi:12793499.
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3341.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgf-bin/cluster.cgi?seq=CS0DD009DD08QPL&cluster=3341.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DD009DD08QPL.
FEATURES
source
location/Qualifiers
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD009YH16"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 242 a 215 c 199 g 225 t
ORIGIN
Query Match 34.48; Score 815.4; D8 9; Length 881;
Best Local Similarity 99.9%; Pred. No. 7e-89;
Matches 816; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 8 GAGCGCTAACGCTCTTTCTGCTCTCCCGCGGTGGTGATGACGGTGAACACTGAGCGTCTA 67
DB 65 GAGCGCTAACGCTCTTTCTGCTCTCCCGCGGTGGTGATGACGGTGAACACTGAGCGTCTA 124
QY 68 AGGCGACCTCATTCTCAGGATGAGGGGCGATGGTGGCAATCTCTATCGCTTTCATGA 127
DB 125 AGGCGACCTCATTCTCAGGATGAGGGGCGATGGTGGCAATCTCTATCGCTTTCATGA 184
QY 128 AGCAGAGGAGGATGGTCTGACGACTTTTATTCAGAGATTGCGCAATACTCTCTATGCAT 187
DB 185 AGCAGAGGAGGATGGTCTGACGACTTTTATTCAGAGATTGCGCAATACTCTCTATGCAT 244

```

```

QY 188 GCAACACCCCTGAAGTTTCAAGTTCATCTTGAAGATCTCCCAACCTCAGGAGCCTGACTTA 247
DB 245 GCAACACCCCTGAAGTTTCAAGTTCATCTTGAAGATCTCCCAACCTCAGGAGCCTGACTTA 304
QY 248 TGAATGCAACCCCTTCTCCTCCAGCAAGTCTTCTCAGCAAAATCAACCTTTGGCCCTCGT 307
DB 305 TGAATGCAACCCCTTCTCCTCCAGCAAGTCTTCTCAGCAAAATCAACCTTTGGCCCTCGT 364
QY 308 CCAATCCTCATGCTTAAACCAATCTGACCTTCTTCTTGAAGTATCGGAAGGCGAGTT 367
DB 365 CCAATCCTCATGCTTAAACCAATCTGACCTTCTTCTTGAAGTATCGGAAGGCGAGTT 424
QY 368 TTGGAAGAGTTCTTCTTACGACACACAAAGCAGAGAGTGTCTTATGCAAGTTTCAAGTTT 427
DB 425 TTGGAAGAGTTCTTCTTACGACACACAAAGCAGAGAGTGTCTTATGCAAGTTTCAAGTTT 484
QY 428 TACAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGAGCATATATATGTCGAGCGGAATG 487
DB 485 TACAGAGAAAGCAATCTCTGAAAGAAAGAGAGAGAGCATATATATGTCGAGCGGAATG 544
QY 488 TTCTGTTGAAGATGTGAAGCACCCCTTTCCTGGTGGGCCCTTCACTTCTTTCACAGCTG 547
DB 545 TTCTGTTGAAGATGTGAAGCACCCCTTTCCTGGTGGGCCCTTCACTTCTTTCACAGCTG 604
QY 548 CTGACAAATTTACTTTTCTCTAGACTACATTAATGTTGAGAGTGTCTTACCATCTCC 607
DB 605 CTGACAAATTTACTTTTCTCTAGACTACATTAATGTTGAGAGTGTCTTACCATCTCC 664
QY 608 AGAGGGAACGCTGCTCTCTGGAACACCGGGCTCTTCTATGCTGCTGGAATAGGAGTGT 667
DB 665 AGAGGGAACGCTGCTCTCTGGAACACCGGGCTCTTCTATGCTGCTGGAATAGGAGTGT 724
QY 668 CCTGGGCTACCTGCTTCACTGACATCGCTTTATAGAGACTTAAACACAGAGAAATTT 727
DB 725 CCTGGGCTACCTGCTTCACTGACATCGCTTTATAGAGACTTAAACACAGAGAAATTT 784
QY 728 TGCTAGATTCACAGGAGACATTTGCTTACTGATTTTCGGACTCTCGAAGAGAGAAATG 787
DB 785 TGCTAGATTCACAGGAGACATTTGCTTACTGATTTTCGGACTCTCGAAGAGAGAAATG 844
QY 788 AACCAACAGCAGACACATCCACCTTCTGTCGACGCC 824
DB 845 AACCAACAGCAGACACATCCACCTTCTGTCGACGCC 881

RESULT 15
LOCUS AL530005/c
DEFINITION AL530005 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD009YH16 3-PRIME, mRNA sequence.
ACCESSION AL530005
VERSION AL530005.1 GI:12793498
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3341.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgf-bin/cluster.cgi?seq=CS0DD009DD08NP1&cluster=3341.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DD009DD08NP1.

```

FEATURES
source Location/Qualifiers
1. 884
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS000003YH16"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 314 a 178 c 146 g 228 t 18 others
ORIGIN

Query Match 34.3%; Score 812.2; DB 9; Length 884;
Best local similarity 95.9%; Pred. No. 1.7e-88;
Matches 848; Conservative 15; Mismatches 18; Indels 3; Gaps 3;

QY 1457 AGCTTAGCAATCTATTGCACACTGTTGCGTGGAAAT-TTTTGAAGAGCACATCTCCCTC 1515
|||||
DB 884 AGCTTAGCAATCTATTGCACACCTGTTGCGTGGAGCKTTTGAAGAGCACATCTCCCTC 825
|||||
QY 1516 AGTGAGCTCATGAGTTTTCATTTTATTCCTCCACCGTGGTCTATCTCTGAAC 1575
|||||
DB 824 AGTGAGCTCATGAGTTTTCATTTTATTCCTCCACCGTGGTCTATCTATGAAC 765
|||||
QY 1576 GAGCGTTAGAGTGGCGCTTAGACGGAGCGAGAGTTTCGTTGAAGAGCGGAC-CTGTTTC 1634
|||||
DB 764 GAGCGTTAGAGTGGCGCTTAGACGGAGCGAGAGTTTCGTTGAAGAGCGGAGCTGTTTC 705
|||||
QY 1635 TAAAGAGGTCTCTGAGAGTCTGTCGGGGTGTGATGACGATATTTATGAATGTC-C 1693
|||||
DB 704 TAAAGAGGTCTCTGAGAGTCTGTCGGGGTGTGATGACGATATTTATGAATGTCGCC 645
|||||
QY 1694 TTTTCTGAAGAGATGTTGTTAGCTCCAAAGCTTTTCTATCGCAGTGTTCAGTCTTTTA 1753
|||||
DB 644 TTTTCTGAAGAGATGTTGTTAGCTCCAAAGCTTTTCTATCGCAGTGTTCAGTCTTTTA 585
|||||
QY 1754 TTTTCCCTTGGGATATGCTGTGTGAACCGCTGCTGAGTGTGTATGCTGATCACAGA 1813
|||||
DB 584 TTTTCCCTTGGGATATGCTGTGTGAACCGCTGCTGAGTGTGTATGCTGATCACAGA 525
|||||
QY 1814 TGGATTTTGTATAGCATCAATGTGACACTTGCAGGACACTACACGCTGGGACATGTT 1873
|||||
DB 524 TGGATTTTGTATAGCATCAATGTGACACTTGCAGGACACTACACGCTGGGACATGTT 465
|||||
QY 1874 TGTTCCTCCATATTTGGAAGATAAATTTATGCTGATGACTTTTGTGAAGATCGGTTAA 1933
|||||
DB 464 TGTTCCTCCATATTTGGAAGATAAATTTATGCTGATGACTTTTGTGAAGATCGGTTAA 405
|||||
QY 1934 TAACTAAATTTATTTGAATGGTCTTGCATGACTCGTATTCAGATGCTTAAAGAAAGCA 1993
|||||
DB 404 TAACTAAATTTATTTGAATGGTCTTGCATGACTCGTATTCAGATGCTTAAAGAAAGCA 345
|||||
QY 1994 TTGCTGCTACAAATATTTCTATTTTGAAGAGGTTTATGACCAATGCCCGAGTTGT 2053
|||||
DB 344 TTGCTGCTACAAATATTTCTATTTTGAAGAGGTTTATGACCAATGCCCGAGTTGT 285
|||||
QY 2054 CAGTCAGACCGCTGGTGTGTTTCATTTTGAATGTCACCTGCTTAAATGGGCATTTAT 2113
|||||
DB 284 CAGTCAGACCGCTGGTGTGTTTCATTTTGAATGTCACCTGCTTAAATGGGCATTTAT 225
|||||
QY 2114 TATGTTTTTTTTTTCATCTCTGATTAATTTGATGATTTGATTAAGAACGCTCTGTACAT 2173
|||||
DB 224 TATGTTTTTTTTTTCATCTCTGATTAATTTGATGATTTGATTAAGAACGCTCTGTACAT 165
|||||
QY 2174 TGGGTTTAAACACTAGTATTTAACTTACAGGCTTATTTGATGATTAACACCATTT 2233
|||||
DB 164 TGGGTTTAAACACTAGTATTTAACTTACAGGCTTATTTGATGATTAACACCATTT 105
|||||
QY 2234 TAAATGACTGTTAATTAACATGTTTATTAATAGCTCAATCTTCCTCATCCCATCACACA 2293
|||||

Db 104 TAAATGACTGTAATTAACATGGTTTATAATAGTACABCCCTCCCTTATGCAATACAC 45
QY 2294 ACTTTTTTTTGTGTGATAAACTGATTTTGTGTTTGCATAAAAC 2337
|||||
Db 44 ACTTTTTTTTGTGTGATAAACTGATTTTGTGTTTGCATAAANC 1

Search completed: August 11, 2003, 06:08:06
Job time : 4498 secs

Sequence 56, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 15, Appli
Sequence 68, Appli
Sequence 39, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 14, Appli
Sequence 1, Appli
Sequence 1433, Ap
Sequence 162, App
Sequence 163, App
Sequence 2, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 25, Appli

28 183 7.7 183 4 US-09-606-421B-56
29 181.8 7.7 1008 4 US-09-394-455-3
30 181.8 7.7 2549 3 US-09-467-082-3
31 181.8 7.7 2549 4 US-09-394-455-5
32 181.8 7.7 2608 4 US-09-394-455-35
33 180.2 7.6 1788 4 US-09-417-197-68
34 180.2 7.6 2211 4 US-09-394-455-39
35 179.2 7.6 2146 4 US-09-842-307-1
36 177.6 7.5 1637 2 US-08-966-316-10
37 175.4 7.4 1619 4 US-09-394-455-14
38 175 7.4 2196 1 US-08-313-274-1
39 189.8 7.2 1636 4 US-09-016-434-1433
40 158.6 6.7 243 1 US-08-594-031-162
41 158.6 6.7 243 1 US-08-594-031-163
42 138.2 5.8 1890 3 US-09-289-466-2
43 136.6 5.8 1891 3 US-09-289-466-1
44 136.6 5.8 1929 2 US-09-016-000-10
45 129.2 5.5 1276 1 US-07-688-352C-25

ALIGNMENTS

RESULT 1
US-09-031-295-1
Sequence 1, Application US/09031295
Patent No. 6326181
GENERAL INFORMATION:
APPLICANT: LANG, Florian
TITLE OF INVENTION: CELL VOLUME-REGULATED HUMAN KINASE H-SGK
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/031.295
FILING DATE: 26-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 197-08-173.8
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 058315/0123
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2370 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 43..1335
US-09-031-295-1

Query Match 100.0%; Score 2370; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2370; Conservative 0; Mismatches 0; Indels 0; Caps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: August 11, 2003, 02:17:47; Search time 161 Seconds
(without alignments)
6497.379 Million cell updates/sec

Title: -
Perfect score: 2370
Sequence: 1 CAGGAGGAGCGCTACGTC.....AAAAAAAAAAAAAAAAAAAA 2370

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 56978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued_Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2370	100.0	2370	4	US-09-031-295-1
2	2262.8	95.5	2311	2	US-08-712-709-6
3	2262.8	95.5	2311	3	US-09-111-444-6
4	2262.8	95.5	2311	3	US-09-541-228-6
5	2262.8	95.5	2311	4	US-09-016-434-772
6	252	10.6	2184	4	US-09-417-197-70
7	252	10.6	2184	4	US-09-417-197-138
8	252	10.6	2610	2	US-09-212-771-1
9	252	10.6	2610	3	US-09-091-058-1
10	245.2	10.3	1599	3	US-09-256-465-1
11	245.2	10.3	1599	4	US-09-167-322-3
12	239.4	10.1	2245	4	US-09-225-749-24
13	232.6	9.8	2274	4	US-09-772-647-3
14	232.2	9.8	2599	6	5266464-1
15	229	9.7	2244	3	US-09-094-714A-48
16	227.6	9.6	2104	3	US-09-313-930-1
17	217	9.2	2751	4	US-09-417-197-72
18	214.6	9.1	2754	4	US-09-429-322-3
19	197	8.3	1732	4	US-09-430-564-1
20	195.2	8.2	2556	4	US-09-817-310-1
21	185.6	7.8	266	1	US-08-594-031-164
22	184.4	7.8	3255	4	US-09-016-434-1471
23	183	7.7	183	3	US-09-040-984-56
24	183	7.7	183	4	US-09-123-912-56
25	183	7.7	183	4	US-09-643-597-56
26	183	7.7	183	4	US-09-480-884A-56
27	183	7.7	183	4	US-09-542-615A-56

1 CACGAGGAGCGCTAAGCTCTTTCTGTCCTCCCGCGGTGGTGGATGACGGTGAACACTGAG 60
1 CACGAGGAGCGCTAAGCTCTTTCTGTCCTCCCGCGGTGGTGGATGACGGTGAACACTGAG 60
61 GCTGCTAAGGCGACCTCACTTACTCAGGATGAGGGCATGGTGGCAATTCATCGCT 120
61 GCTGCTAAGGCGACCTCACTTACTCAGGATGAGGGCATGGTGGCAATTCATCGCT 120
121 TTCATGAAGCAGAGGATGGGTCTCAACGACTTTATTCCAGAAATTCGCCAATACTCC 180
121 TTCATGAAGCAGAGGATGGGTCTCAACGACTTTATTCCAGAAATTCGCCAATACTCC 180
181 TATGATGCAACACCTCGAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAGCCT 240
181 TATGATGCAACACCTCGAAGTTCAGTCCATCTTGAAGATCTCCCAACCTCAGAGCCT 240
241 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCAGCAAAATCAACCTTGGC 300
241 GAGCTTATGAATGCCAACCCCTTCTCTCCACCAAGTCTCTCAGCAAAATCAACCTTGGC 300
301 CGTGCCTCAATCTCATGCTTAACCATCTGACTTTCACCTCTTGAAGTGAATCGGAAG 360
301 CGTGCCTCAATCTCATGCTTAACCATCTGACTTTCACCTCTTGAAGTGAATCGGAAG 360
361 GGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAGGAGGAGGAGGAGGAGGAGGAGGAG 420
361 GGCAGTTTGGAAAGGTTCTTCTAGCAAGACACAGGAGGAGGAGGAGGAGGAGGAGGAG 420
421 AAAGTTTACAGAGAAAGCAATCTCGAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
421 AAAGTTTACAGAGAAAGCAATCTCGAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
481 CGAATGTCTGTGGAAGATGTGAAGCAGCCTTCTCTGTTGGGCTTCACTTCTCTTTC 540
481 CGAATGTCTGTGGAAGATGTGAAGCAGCCTTCTCTGTTGGGCTTCACTTCTCTTTC 540
541 CAGACTGCTGACAAATCTACTTGTCTAGACTACATTAATGTTGAGAGGAGTGTCTTC 600
541 CAGACTGCTGACAAATCTACTTGTCTAGACTACATTAATGTTGAGAGGAGTGTCTTC 600
601 CATCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
601 CATCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
661 GCGAGTCCCTTGGGCTACCTGCATCTGACTGACATCTTTATAGAGACTTAACACAGAG 720
661 GCGAGTCCCTTGGGCTACCTGCATCTGACTGACATCTTTATAGAGACTTAACACAGAG 720
721 AATATTTGCTAGATTCACAGGACACATCTGCTTACTGATTTCCGACTCTGCAAGGAG 780
721 AATATTTGCTAGATTCACAGGACACATCTGCTTACTGATTTCCGACTCTGCAAGGAG 780
781 ACATTTGAACACACAGGACACATCTGCTTACTGATTTCCGACTCTGCAAGGAG 840
781 ACATTTGAACACACAGGACACATCTGCTTACTGATTTCCGACTCTGCAAGGAG 840
841 GAGGTGCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
841 GAGGTGCTTATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
901 TTGTATGAGATCTGATGGCTGCGCTTTTATAGCGGAAACACAGCTGAAATGTAC 960
901 TTGTATGAGATCTGATGGCTGCGCTTTTATAGCGGAAACACAGCTGAAATGTAC 960
961 GACAACTTCTGAACAAAGCCTCTCCAGCTGAACCAATATTAACAAATCCGCAAGACAC 1020
961 GACAACTTCTGAACAAAGCCTCTCCAGCTGAACCAATATTAACAAATCCGCAAGACAC 1020
1021 CTCCTGGAGGCTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
1021 CTCCTGGAGGCTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080

1081 ATGAGATTAAGAGTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
1081 ATGAGATTAAGAGTCAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
1141 AAGATTACTCCCT 1200
1141 AAGATTACTCCCT 1200
1201 CCCGAGTTTACGAGAGGCTTCT 1260
1201 CCCGAGTTTACGAGAGGCTTCT 1260
1261 GTCACAGCCGCTCAAGAGAGCTGCGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
1261 GTCACAGCCGCTCAAGAGAGCTGCGAGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
1321 ACGGACTCTTCT 1380
1321 ACGGACTCTTCT 1380
1381 CGAATGTTTATGTTAGCTTCT 1440
1381 CGAATGTTTATGTTAGCTTCT 1440
1441 TTTGCACATCTCTGGAAGCTTACCAATCTTATTGCACACTGTTTCGCTGGAATTTTGA 1500
1441 TTTGCACATCTCTGGAAGCTTACCAATCTTATTGCACACTGTTTCGCTGGAATTTTGA 1500
1501 GAGCAGATCT 1560
1501 GAGCAGATCT 1560
1561 TGCTATCTCTGAAACAGCGTTAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
1561 TGCTATCTCTGAAACAGCGTTAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1620
1621 AAGCGGACT 1680
1621 AAGCGGACT 1680
1681 TATGAAATGTGCTTCT 1740
1681 TATGAAATGTGCTTCT 1740
1741 TTTGAGTCT 1800
1741 TTTGAGTCT 1800
1801 GCTGATCAGATGGAATTTTCT 1860
1801 GCTGATCAGATGGAATTTTCT 1860
1861 GTGGGACATTTGTTGTTCT 1920
1861 GTGGGACATTTGTTGTTCT 1920
1921 AAGATACCGTCT 1980
1921 AAGATACCGTCT 1980
1981 CCTAAGAAAGCATTGCTGCTTCT 2040
1981 CCTAAGAAAGCATTGCTGCTTCT 2040
2041 ATGCCCCAGTCTCAGTCAAGCGCTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2041 ATGCCCCAGTCTCAGTCAAGCGCTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2101 ATGCGGCTATTTATCT 2160
2101 ATGCGGCTATTTATCT 2160
2161 AACGCTCTGATCTTGGGTTTATACACTAGTATATTTAACTTACAGGCTTATTTGTAAATG 2220

Db 2161 AACGCTGTACATGGGTTAACAACCTAGTATATTTAACTTACAGGCTTATTTGTAATG 2220
Qy 2221 TAAACCAACCTTTTAAATGTACTCTAATTAACATGGTATATAATACGTPACAACTCTTCCCTC 2280
Db 2221 TAAACCAACCTTTTAAATGTACTCTAATTAACATGGTATATAATACGTPACAACTCTTCCCTC 2280
Qy 2281 ATCCATCATCACACACTTTTTCGTGTGTGATGAACATGATTTTGGTTTGGCAATAAACAACCTT 2340
Db 2281 ATCCATCATCACACACTTTTTCGTGTGTGATGAACATGATTTTGGTTTGGCAATAAACAACCTT 2340
Qy 2341 GAAAAATAAAAAAATAAAAAAATAAAAAA 2370
Db 2341 GAAAAATAAAAAAATAAAAAAATAAAAAA 2370

RESULT 2
US-08-712-709-6
; Sequence 6, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PG-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
US-08-712-709-6

Query Match 95.5%; Score 2262.8; DB 2; Length 2311;
Best local Similarity 99.5%; Pred. No. 0;
Matches 2300; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
Qy 34 GCGGTGGTGTACAGCGTGAACACTGAGCTGCTAAGGCACCCCTCACTTACTCCAGATG 93
Db 1 GCGGTGGTGTACAGCGTGAACACTGAGCTGCTAAGGCACCCCTCACTTACTCCAGATG 60
Qy 94 AGGGCATGGTGGCAATCTCATCGCTTTTCATGAAGCAGAGGAGGAGTGGTCTGAACAC 153
Db 61 AGGGCATGGTGGCAATCTCATCGCTTTTCATGAAGCAGAGGAGGAGTGGTCTGAACAC 120
Qy 154 TTTATTACAGAAATGCCAATAACTCTTATGCATGCAACACACCTGAATTCAGTCCATC 213

Db 121 TTTATTACAGAAATGCCAATAACTCTTATGCATGCAACACACCTGAATTCAGTCCATC 180
Qy 214 TTGAAGATCTCCCAACCTCAGGAGCTGAGCTTATGAATGCCAACCTTCTCTCCACCA 273
Db 181 TTGAAGATCTCCCAACCTCAGGAGCTGAGCTTATGAATGCCAACCTTCTCTCCACCA 240
Qy 274 AGTCTCTCTCAGCAAAATCAACCTTTGGCCCGCTGCTCCAAATCTCATCTAAACCAATCAG 333
Db 241 AGTCTCTCTCAGCAAAATCAACCTTTGGCCCGCTGCTCCAAATCTCATCTAAACCAATCAG 300
Qy 334 TTTCACCTCTTGAAGATGTCGGAAGGGCAGCTTTTGGGAAGGCTTCTTCTAGCAACACAC 393
Db 301 TTTCACCTCTTGAAGATGTCGGAAGGGCAGCTTTTGGGAAGGCTTCTTCTAGCAACACAC 360
Qy 394 AAGCAGCAAGATGTTCTTATGCACTCAAAATTTTACAGAAAGAAACAAATCTTGAAGAA 453
Db 361 AAGCAGCAAGATGTTCTTATGCACTCAAAATTTTACAGAAAGAAACAAATCTTGAAGAA 420
Qy 454 AAGAGGAGAAAGCATATATATGTCGGAGCGGAATGTTCTGTTGAAGATGTGAAGCACTCT 513
Db 421 AAGAGGAGAAAGCATATATATGTCGGAGCGGAATGTTCTGTTGAAGATGTGAAGCACTCT 480
Qy 514 TTCTGTGTGGGCTTCACTTCTTTCAGACTGCTCACAATTTGTACTTTGTCTGAGAC 573
Db 481 TTCTGTGTGGGCTTCACTTCTTTCAGACTGCTCACAATTTGTACTTTGTCTGAGAC 540
Qy 574 TACATTAATGTGGAGAGTGTCTTCTACCATCTTCAGAGGAAACGCTGCTTCTGGAACCA 633
Db 541 TACATTAATGTGGAGAGTGTCTTCTACCATCTTCAGAGGAAACGCTGCTTCTGGAACCA 600
Qy 634 CGGGCTCTTCTTATGCTGCTGAATAGCCAGTCCCTTGGCTACCTGCCATTCATCTGAC 693
Db 601 CGGGCTCTTCTTATGCTGCTGAATAGCCAGTCCCTTGGCTACCTGCCATTCATCTGAC 660
Qy 694 ATCGTTTATAGAGACTTAAACACACAGAAATTTTGTCTAGATTCACAGGAGCAATTTGTC 753
Db 661 ATCGTTTATAGAGACTTAAACACACAGAAATTTTGTCTAGATTCACAGGAGCAATTTGTC 720
Qy 754 CTACTGATTTTGGACTCTGCAAGGAGAAATTTGAACACAGACACACACATCTCAATCTTC 813
Db 721 CTACTGATTTTGGACTCTGCAAGGAGAAATTTGAACACAGACACACACATCTCAATCTTC 780
Qy 814 TGTGCACGCGGAGTATCTCGACCTGAGGTGCTTATGATGAGATGCTGATGGCTTGGCTTTT 933
Db 781 TGTGCACGCGGAGTATCTCGACCTGAGGTGCTTATGATGAGATGCTGATGGCTTGGCTTTT 900
Qy 841 GTGGACTGGTGGCTGGGAGTGTCTTGTATGAGATGCTGATGGCTTGGCTTTT 900
Db 934 TATAGCGGAACACAGCTGAAATGTACGACAACTTCTGAAACAGCTTCTGAGCTGAAA 993
Qy 901 TATAGCGGAACACAGCTGAAATGTACGACAACTTCTGAAACAGCTTCTGAGCTGAAA 960
Qy 994 CCAAAATATTACAAATTCGCAAGACACCTCTGAGGCGCTTCTGCAAGAGGAGGAGCA 1053
Db 961 CCAAAATATTACAAATTCGCAAGACACCTCTGAGGCGCTTCTGCAAGAGGAGGAGCA 1020
Qy 1054 AAGCGCTCTGGGCGCAAGGATGATCTCATGAAGATTAAGAGTCTATGCTTCTTCTCTCTA 1113
Db 1021 AAGCGCTCTGGGCGCAAGGATGATCTCATGAAGATTAAGAGTCTATGCTTCTTCTCTA 1080
Qy 1114 ATTACTGGGATGATCTCATTAATGAAGATTAAGAGTCTATGCTTCTTCTCTCTA 1173
Db 1081 ATTACTGGGATGATCTCATTAATGAAGATTAAGAGTCTATGCTTCTTCTCTCTA 1140
Qy 1174 GGGCCCAACAGGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGGAGCTTCTCCCACTCC 1233
Db 1141 GGGCCCAACAGGAGCTACGGCACTTTGACCCCGAGTTTACCGAAGAGGAGCTTCTCCCACTCC 1200
Qy 1234 ATTGCAAGTCCCTTGACAGGCTCTCTGTCACAGGAGGCTCAAGAGGAGCTGCGGAGCT 1293

Db 1201 ATTGCCAAGTCCCTGACAGGGTCTCTGTCACAGGCGAGCTCAAGGAAGCTGCGGAGGCT 1260
 QY 1294 TTCCTAGGCTTTCTCTAGGCGCTCCACCGAGCTCTTCTCTGTAACCTCTTAGGGCTT 1353
 Db 1261 TTCCTAGGCTTTCTCTAGGCGCTCCACCGAGCTCTTCTCTGTAACCTCTTAGGGCTT 1320
 QY 1354 GGTTTAAAGGATTTATGTGTGTTCCGGAATGTTTASGTAGCTTTTGGTGGAGCCGC 1413
 Db 1321 GGTTTAAAGGATTTATGTGTGTTCCGGAATGTTTASGTAGCTTTTGGTGGAGCCGC 1380
 QY 1414 CAGCTGACAGACATCTACAGAGAATTTGCACATCTCGAAGCTTAGCATCTATT 1473
 Db 1381 CAGCTGACAGACATCTACAGAGAATTTGCACATCTCGAAGCTTAGCATCTATT 1440
 QY 1474 GCACACTGTCCTCGTGAAT-TTTTTTGAAGAGCACATCTCTCAGTCAGCTCATGAGT 1532
 Db 1441 GCACACTGTCCTCGTGAAGCTTTTGAAGAGCACATCTCTCAGTCAGCTCATGAGT 1500
 QY 1533 TTCATTTTATCTCTCTCAAGCTGGTCTATCTGTAAGAGCGGTAGAGTCCGC 1592
 Db 1501 TTCATTTTATCTCTCTCAAGCTGGTCTATCTGTAAGAGCGGTAGAGTCCGC 1560
 QY 1593 CTTAGACGAGGAGGAGTTCGTAGAAAGCGAC-CTCTCTTAAAGAGGCTCTCTGC 1651
 Db 1561 CTTAGACGAGGAGGAGTTCGTAGAAAGCGAGCTCTCTTAAAGAGGCTCTCTGC 1620
 QY 1652 AGATCTCTCTGGCTGTCATCAGCAATATATGAATGTCGCTTTCTGAAG-AGATTT 1710
 Db 1621 AGATCTCTCTGGCTGTCATCAGCAATATATGAATGTCGCTTTCTGAAG-AGATTT 1680
 QY 1711 GTTAGCTCAAGCTTTTCTCTATCAGTGTTCAGTCTCTTATTTTCCCTTGTGGAT 1770
 Db 1681 GTTAGCTCAAGCTTTTCTCTATCAGTGTTCAGTCTCTTATTTTCCCTTGTGGAT 1740
 QY 1771 GCTGTGTAAGCTGCTGTGAGTGTGATGCTGATCAGATGGAATTTGTTATAAGC 1830
 Db 1741 GCTGTGTAAGCTGCTGTGAGTGTGATGCTGATCAGATGGAATTTGTTATAAGC 1800
 QY 1831 ATCAATGTGACACTTTCGAGGACACTACAGCTGGGACATGTTGTTTCTTCCATTTG 1890
 Db 1801 ATCAATGTGACACTTTCGAGGACACTACAGCTGGGACATGTTGTTTCTTCCATTTG 1860
 QY 1891 GAAGATAAATTTATGTGTAGACTTTTGTGAAGTACGGTTTAACTAAATTTATTGA 1950
 Db 1861 GAAGATAAATTTATGTGTAGACTTTTGTGAAGTACGGTTTAACTAAATTTATTGA 1920
 QY 1951 ATGCTCTTGCATCTCTATTCAGATGCTTAAAGAACCATTCCTCTACAAATATT 2010
 Db 1921 ATGCTCTTGCATCTCTATTCAGATGCTTAAAGAACCATTCCTCTACAAATATT 1980
 QY 2011 TCTATTTTGAAGAGGTTTATGACCAATGCCCCAGTTGTGTCAGTCAGAGCGGTTGGT 2070
 Db 1981 TCTATTTTGAAGAGGTTTATGACCAATGCCCCAGTTGTGTCAGTCAGAGCGGTTGGT 2040
 QY 2071 GTTTTCTATGTTTAAATGTCNCTGTAAATGGGATTTATGTTTCTTCTTCTTTCG 2130
 Db 2041 GTTTTCTATGTTTAAATGTCNCTGTAAATGGGATTTATGTTTCTTCTTCTTTCG 2100
 QY 2131 ATTCCTGATAATTTGATGTTATGTAAGAACGCTCTGATACATCGGTTATACACTAGT 2190
 Db 2101 ATTCCTGATAATTTGATGTTATGTAAGAACGCTCTGATACATCGGTTATACACTAGT 2160
 QY 2191 ATATTTAACCTTACAGGCTTATTTGTAATGTAAGAACCATTTTAAATGCTACTGTAATA 2250
 Db 2161 ATATTTAACCTTACAGGCTTATTTGTAATGTAAGAACCATTTTAAATGCTACTGTAATA 2220
 QY 2251 CATGGTTATAATACGTAATCTTCCCTCATCCCATCACACAACCTTTTGTGTGTGA 2310
 Db 2221 CATGGTTATAATACGTAATCTTCCCTCATCCCATCACACAACCTTTTGTGTGTGA 2280
 QY 2311 TAACTGATTTTGGTTTGAATAAAGCTTG 2341
 Db 2281 TAACTGATTTTGGTTTGAATAAAGCTTG 2311

RESULT 3
 US-09-111-444-6
 ; Sequence 6, Application US/09111444
 ; Patent No. 6045792
 ; GENERAL INFORMATION:
 ; APPLICANT: Au-Young, Janice
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Hawkins, Phillip R.
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: U.S.
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/111.444
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/712,709
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0118 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2311 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; IMMEDIATE SOURCE:
 ; LIBRARY:
 ; CLONE: Consensus
 ; US-09-111-444-6

Query Match 95.5%; Score 2262.8; DB 3; Length 2311;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2300; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
 QY 34 CGGCTGGTGTGATGAGCTGAAACCTGAGGCTGCTAAGGGCACCTCTACTTACTTCCAGGATG 93
 Db 1 CGGCTGGTGTGATGAGCTGAAACCTGAGGCTGCTAAGGGCACCTCTACTTACTTCCAGGATG 60
 QY 94 AGGGCATGGTGGCAATTTCTATCGCTTTCATGAAGCAGAGGAGGATGGTCTGTAACGAC 153
 Db 61 AGGGCATGGTGGCAATTTCTATCGCTTTCATGAAGCAGAGGAGGATGGTCTGTAACGAC 120
 QY 154 TTTATTCAGAGAGTGGCAATTTCTATCGCTTTCATGAAGCAGAGGAGGATGGTCTGTAACGAC 213
 Db 121 TTTATTCAGAGAGTGGCAATTTCTATCGCTTTCATGAAGCAGAGGAGGATGGTCTGTAACGAC 180
 QY 214 TTGAAGATCTCCCAACCTCAGAGGAGGATGGTCTGTAACGAGGATGGTCTGTAACGAC 273
 Db 181 TTGAAGATCTCCCAACCTCAGAGGAGGATGGTCTGTAACGAGGATGGTCTGTAACGAC 240
 QY 274 AGTCTTCTCAGCAAAATCAACCTTGGCGCGCTGCTGCTCCATCTCTATCTTAACCATCTGAC 333
 Db 241 AGTCTTCTCAGCAAAATCAACCTTGGCGCGCTGCTGCTCCATCTCTATCTTAACCATCTGAC 300

334 TTTCACTTCTTGAAGTATCGGAAAGGCGAGTTTGGAAAGGTTCTTCTAGCAAGACAC 393
301 TTTCACTTCTTGAAGTATCGGAAAGGCGAGTTTGGAAAGGTTCTTCTAGCAAGACAC 360
394 AAGCGAGAAGAGTGTCTATGCAAGTCAAAAGTTTACAGAGAAAGCAATCCCGAAGAAAG 453
361 AAGCGAGAAGAGTGTCTATGCAAGTCAAAAGTTTACAGAGAAAGCAATCCCGAAGAAAG 420
454 AAGAGGAGAGAGTATATGTCGAGCGGAATGTTCTGTTGAAGATGTAAGAGCAACCT 513
421 AAGAGGAGAGAGTATATGTCGAGCGGAATGTTCTGTTGAAGATGTAAGAGCAACCT 480
514 TTCCTGGTGGGCTTCACTTCTTCCAGAGTCTGCAAAATGTAAGTCTTCTAGAC 573
481 TTCCTGGTGGGCTTCACTTCTTCCAGAGTCTGCAAAATGTAAGTCTTCTAGAC 540
574 TACATTAATGTCGAGAGTCTTCTTACCATCTCCAGAGGAGCACTGCTTCTCCGGAACCA 633
541 TACATTAATGTCGAGAGTCTTCTTACCATCTCCAGAGGAGCACTGCTTCTCCGGAACCA 600
634 CGGGCTGCTTCTATGTCGTCGAAATAGCCAGTGCCTTGGGCTTACCTGCAATCACTGAAC 693
601 CGGGCTGCTTCTATGTCGTCGAAATAGCCAGTGCCTTGGGCTTACCTGCAATCACTGAAC 660
694 ATCGTTTATAGAGACTTAAACACAGAGAAATATTTGCTAGATTACAGAGGAGCAATGTC 753
661 ATCGTTTATAGAGACTTAAACACAGAGAAATATTTGCTAGATTACAGAGGAGCAATGTC 720
754 CTTACTGATTTGCGAGCTGCAAGAGAGAAATGTAACACACACAGAGCAATCCACCTTC 813
721 CTTACTGATTTGCGAGCTGCAAGAGAGAAATGTAACACACACAGAGCAATCCACCTTC 780
814 TGTGACACCGCCGAGTATCTCGACCTGAGTGTCTATAGAGAGCTTATGACAGGACT 873
781 TGTGACACCGCCGAGTATCTCGACCTGAGTGTCTATAGAGAGCTTATGACAGGACT 840
874 GTGGAGTGTGTGCTGGAGCTGTCTTGTATGAGATGCTGTATGGCTTCCGCCCTTT 933
841 GTGGAGTGTGTGCTGGAGCTGTCTTGTATGAGATGCTGTATGGCTTCCGCCCTTT 900
934 TATAGCGAAACACAGCTGAATGTACGACAACTTCTGAACAAAGCTTCCAGCTGAAA 993
901 TATAGCGAAACACAGCTGAATGTACGACAACTTCTGAACAAAGCTTCCAGCTGAAA 960
994 CCAATATATACAAATTCGCAAGACACCTCTCGAGGCTTCCGAGAGGAGGAGCA 1053
961 CCAATATATACAAATTCGCAAGACACCTCTCGAGGCTTCCGAGAGGAGGAGCA 1020
1054 AAGCGCTCGGGCCCAAGGATGACTTCTATGAGATTAAGAGTCTGCTTCTCTCTTA 1113
1021 AAGCGCTCGGGCCCAAGGATGACTTCTATGAGATTAAGAGTCTGCTTCTCTCTTA 1080
1114 ATTAAGTGGATGATCTAATTAAGAGATTAAGAGTCTGCTTCTCTCTTA 1173
1081 ATTAAGTGGATGATCTAATTAAGAGATTAAGAGTCTGCTTCTCTCTTA 1140
1174 GGCCCAAGAGCTAGGCACTTGAACCCGAGTTTACGAGAGGCTTGTCCCAACTCC 1233
1141 GGCCCAAGAGCTAGGCACTTGAACCCGAGTTTACGAGAGGCTTGTCCCAACTCC 1200
1234 ATTGCAAGTCTCCGAGAGCTCTCTGTCAGAGGCTTCAAGAGGCTTGGCGAGGCT 1293
1201 ATTGCAAGTCTCCGAGAGCTCTCTGTCAGAGGCTTCAAGAGGCTTGGCGAGGCT 1260
1294 TTCTAGGCTTTTCTATGCGCTCCAGGAGCTTCTTCTCTGAAACCTGTAGGCGCT 1353
1261 TTCTAGGCTTTTCTATGCGCTCCAGGAGCTTCTTCTCTGAAACCTGTAGGCGCT 1320
1354 GGTATTAAGAGATTTATGCTGCTTCCGATGTTTATGCTAGGCTTTTGGTGGAGCGC 1413
1321 GGTATTAAGAGATTTATGCTGCTTCCGATGTTTATGCTAGGCTTTTGGTGGAGCGC 1380
1414 CAGCTGACAGGACATCTTACAGAGAAATTTGACATCTCTGGAAGCTTATGCAATCTTAT 1473

1381 CAGCTGACAGAGACATCTTACAGAGAAATTTGCAATCTCTGAGGCTTAGCAATCTTAT 1440
1474 GCACACTCTTCTGCTGGAATTTTGTGAAGAGACATCTCTCAGTCAAGTCAAGAGT 1532
1441 GCACACTCTTCTGCTGGAAGCTTTTGTGAAGAGACATCTCTCAGTCAAGTCAAGAGT 1500
1533 TTCATTTATTTCTTCTTCCAACTGCTATCTCTGAAACGAGCGTTAGAGTCTGCTC 1592
1501 TTCATTTATTTCTTCTTCCAACTGCTATCTCTGAAACGAGCGTTAGAGTCTGCTC 1560
1593 CTTAGACGAGGAGGAGGCTTCTGTTAGAAAGCGGAC-CTGTTCTTAAAGAGTCTGCTC 1651
1561 CTTAGACGAGGAGGAGGCTTCTGTTAGAAAGCGGAGCTGTTCTTAAAGAGTCTGCTC 1620
1652 AGATCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1710
1621 AGATCTGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1711 GTTACTCTCAAGCTTTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1770
1681 GTTACTCTCAAGCTTTTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
1771 GCTGTTGAACCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1830
1741 GCTGTTGAACCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1831 ATCAATGTCACACTTGCAGGAGCTTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890
1801 ATCAATGTCACACTTGCAGGAGCTTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1891 GAAGATAAATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1950
1861 GAAGATAAATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1951 AATGCTGCTGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2010
1921 AATGCTGCTGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
2011 TCTATTTTGAAGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2070
1981 TCTATTTTGAAGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2071 GTTATTTTGAAGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2130
2041 GTTATTTTGAAGAGGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2131 ATTCCTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2190
2101 ATTCCTGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2191 ATATTTAAACTTACAGGCTTATTTCTGATGATGATGATGATGATGATGATGATGATGAT 2250
2161 ATATTTAAACTTACAGGCTTATTTCTGATGATGATGATGATGATGATGATGATGATGAT 2220
2251 CATGTTTATATACGTAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2310
2221 CATGTTTATATACGTAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
2311 TAACTGATTTTGGTTTGGTAAATTAACCTTG 2341
2281 TAACTGATTTTGGTTTGGTAAATTAACCTTG 2311

RESULT 4
US-09-541-228-6
; Sequence 6, Application US/09541228
; Patent No. 6232077
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES


```

QY 1593 CTTAGACGAGGAGGAGTTTCCTTAGAAAGCGAC-CGTGTTCTAAAAAGTCTCCTGC 1651
DB 1561 CTTAGACGAGGAGGAGTTTCCTTAGAAAGCGAGCTGTTCTAAAAAGGTCCTCCTGC 1620
QY 1652 AGATCTCTCTGGGCTGTGATGACGAATATTAAGAAATGTCCTTCTCTGAAG-AGATGCT 1710
DB 1621 AGATCTCTCTGGGCTGTGATGACGAATATTAAGAAATGTCCTTCTCTGAAGAAATGCT 1680
QY 1711 GTTAGCTCCAAAGCTTTTCTCTATCGCAGTGTTCAGTCTCTTATTTTCTCTGTGATAT 1770
DB 1681 GTTAGCTCCAAAGCTTTTCTCTATCGCAGTGTTCAGTCTCTTATTTTCTCTGTGATAT 1740
QY 1771 CTTGTGTGAACCGTCTGTGATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 1830
DB 1741 CTTGTGTGAACCGTCTGTGATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 1800
QY 1831 ATCAATGTGACACTTTCAGGACACTACACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1890
DB 1801 ATCAATGTGACACTTTCAGGACACTACACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1860
QY 1891 GAAGATAAATTTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 1950
DB 1861 GAAGATAAATTTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 1920
QY 1951 AATGCTCTTGAACGACTTTCAGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 2010
DB 1921 AATGCTCTTGAACGACTTTCAGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 1980
QY 2011 TCTATTTTGAAGAGGTTTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 2070
DB 1981 TCTATTTTGAAGAGGTTTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTC 2040
QY 2071 GTTTTTCATCTTTTAAATGTCACCTGTAATGTCGTATGTCGTATGTCGTATGTCGTATG 2130
DB 2041 GTTTTTCATCTTTTAAATGTCACCTGTAATGTCGTATGTCGTATGTCGTATGTCGTATG 2100
QY 2131 ATCTCTGATATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 2190
DB 2101 ATCTCTGATATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 2160
QY 2191 ATATTAAACTTACAGGCTTATTTGTAATGTCGTATGTCGTATGTCGTATGTCGTATGTC 2250
DB 2161 ATATTAAACTTACAGGCTTATTTGTAATGTCGTATGTCGTATGTCGTATGTCGTATGTC 2220
QY 2251 CATGCTTATATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 2310
DB 2221 CATGCTTATATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGTATGTCGT 2280
QY 2311 TAACTGATTTTGGTTTCCATAAAACCTTG 2341
DB 2281 TAACTGATTTTGGTTTCCATAAAACCTTG 2311

```

RESULT 5

```

US-09-016-434-772
; Sequence 772, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 772:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2311 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: MMLR2DT01
; CLONE: 477245
; US-09-016-434-772

```

```

Query Match 95.58; Score 2262.8; DB 4; Length 2311;
Best local Similarity 99.58; Pred. No. 0;
Matches 2300; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 34 GCGGTGCTGATGACGGTGAAACCTGAGGCTGCTAAGGCGACCTCACTTACTCCAGATG 93
DB 1 GCGGTGCTGATGACGGTGAAACCTGAGGCTGCTAAGGCGACCTCACTTACTCCAGATG 60
QY 94 AGGGCGTGTGGCAATTCATCGCTTTCATGAGCAGAGGAGGATGAGTCAATGAC 153
DB 61 AGGGCGTGTGGCAATTCATCGCTTTCATGAGCAGAGGAGGATGAGTCAATGAC 120
QY 154 TTTATTCAGAGATGTCCTTATGATGTCATGTCATGTCATGTCATGTCATGTCATGTC 213
DB 121 TTTATTCAGAGATGTCCTTATGATGTCATGTCATGTCATGTCATGTCATGTCATGTC 180
QY 214 TTGAGATGTCCTCCAACTTCAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 273
DB 181 TTGAGATGTCCTCCAACTTCAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 240
QY 274 AGTCTCTCTCAGCAATCAACCTTGGGCGCTGCTCCATCTCTCATGTCATGTCATGTC 333
DB 241 AGTCTCTCTCAGCAATCAACCTTGGGCGCTGCTCCATCTCTCATGTCATGTCATGTC 300
QY 334 TTTCTACTTCTCAAGTGCATCGAAAGGCGAGTTTGGAAAGGTTCTTCTAGTAAAGAC 393
DB 301 TTTCTACTTCTCAAGTGCATCGAAAGGCGAGTTTGGAAAGGTTCTTCTAGTAAAGAC 360
QY 394 AAGGCGAGAGAGTGTCTTATGTCAGTCAAAAGTTTACAGAGAAAGCAATCTTCAAAAG 453
DB 361 AAGGCGAGAGAGTGTCTTATGTCAGTCAAAAGTTTACAGAGAAAGCAATCTTCAAAAG 420
QY 454 AAGAGGAGAGAGCATATTTATGTCGAGCGGAGTGTCTGTTGAGATGTCGAGCAGCT 513
DB 421 AAGAGGAGAGAGCATATTTATGTCGAGCGGAGTGTCTGTTGAGATGTCGAGCAGCT 480
QY 514 TTTCTGTTGGGCTTCTACTTCTTCTTCTCAGACTCTGACAAATGTCATTTTCTTCTAG 573
DB 481 TTTCTGTTGGGCTTCTACTTCTTCTTCTCAGACTCTGACAAATGTCATTTTCTTCTAG 540
QY 574 TACATTAATGTCGAGAGTGTCTTCTTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 633
DB 541 TACATTAATGTCGAGAGTGTCTTCTTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 634 GCGGCTCTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 693

```


445 GAGTTTGAGTACCTGAGCTGCTGGCAAGGCGACCTTTCGGCAAGGTGATCCTGGTGAAG 504
QY
391 CACAAGGAGAGAGTGTCTATGCTCAAGTATTTACAGAAGAAAGAACTCCTGAAA 450
Db
505 GAGAGGCGACAGCGGCTACTACCCATGATGATCTCAAGAAGAGAGTATCTGCGCC 564
QY
451 AAGAAGAGGAGAGACATATATGCGGAGCGGAATGTTCTGTGAAGAATGTGAAGCAC 510
Db
565 AAGGACGAGTGGTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 621
QY
511 CTTTCTGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
Db
622 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
QY
571 GACTACATTAATGTTGGAGAGTGTCTACATCTTCAGAGGAGGAGGAGGAGGAGGAGG 630
Db
682 GAGTACGAGCAAGCGGCGGAGTGTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 741
QY
631 CCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
Db
742 GACCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
QY
688 CTGAACATCGTTTATAGAGACTTAAACACAGAGATATTTTGTGATGATTCACAGGACAC 747
Db
802 AAGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
QY
748 ATTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Db
862 ATTAAGTACAGAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
QY
808 ACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
Db
922 ACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981
QY
868 AGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Db
982 CGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
QY
928 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987
Db
1042 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
QY
988 CTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCGCTCTCTGAGAGGAC 1047
Db
1102 TTTCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1161
QY
1048 AGGACAAAGCGCT---CGGGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
Db
1162 CCAAGCAGAGGCTTGGCGGGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
QY
1105 TTTCTCTTAATTAAGTGGATGATCTCAATTAAGAAGATTTACTCCCGCTTTTAACCCA 1164
Db
1222 TTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
QY
1165 AATGTGAGTGGGCGCCAGCAGCTAGGCGACTTTGAGCCCGGAGTTTACCGAGAGGCGCTG 1224
Db
1282 CAGGTACGCTGCGAGACTGACACAGGATTTTGTGAGGAGTTTACCGGCGGCGGAGTAT 1341
QY
1225 CCAAGCTC 1232
Db
1342 ACCATAC 1349

RESULT 7

US-09-417-197-138
; Sequence 138, Application US/09417197
; Patent No. 6518021
; GENERAL INFORMATION:
; APPLICANT: Ole TRASTUP, et al.
; TITLE OF INVENTION: A Method For Extracting Quantitative Information Relating To An
; FILE REFERENCE: 3759-0110P
; CURRENT APPLICATION NUMBER: US/09/417.197

; CURRENT FILING DATE: 1999-10-07
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: patentIn version 3.0
; SEQ ID NO 138
; LENGTH: 2184
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EGFP-PKB fusion
; NAME/KEY: CDS
; LOCATION: (1)..(2181)
; US-09-417-197-138

Query Match 10.6%; Score 252; DB 4; Length 2184;

Best Local Similarity 57.2%; Pred. No. 6e-60;

Matches 519; Conservative 0; Mismatches 380; Indels 9; Gaps 3;

QY 331 GACTTTCACCTTCTTGAAGTATATCGGAAGGCGACCTTTCGAAAAGGTTCTTCTTACAAAGA 390
Db 1186 GAGTTTGAGTACTCTGAGCTGCTGGCAAGGCGACTTTCGCAAGGTGATCTCTGCTGAAG 1245
QY 391 CACAAGGAGAGAGAGTGTCTATGCTCAAGTGTTCACAGAAGAAAGCAATCTTAA 450
Db 1246 GAGAAGGCGCACAGGCGCTACTACGCCATGAGATCTCTCAAGAAGGAAGTCTCTGCTGCT 1305
QY 451 AAGAAGGAGGAGAGACATATATGCTGGAGGAGGATGTTCTGTTGAAGAATGTGAAGAC 510
Db 1306 AAGGACGAGGTGGCCACACACTCACCA---GAACCGGCTCTGACAGACTCCAGGAC 1362
QY 511 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570
Db 1363 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
QY 571 GACTACATTAATGTTGGAGAGTGTCTTACCATCTCCAGAGGAGGAGGAGGAGGAGGAG 630
Db 1423 GAGTACGCGCAAGCGGCGGAGCTGTTCTTCCACTGTCGCGGAGACGTGTTCTGCTGCT 1482
QY 631 CCAGCGGCTGCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 687
Db 1483 GACCGCGCGCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1542
QY 688 CTGAACATCGTTTATAGAGACTTAAACACAGAGATATTTGCTAGATTCACAGAGAC 747
Db 1543 AAGAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1602
QY 748 ATTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Db 1603 ATTAAGTACAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1662
QY 808 ACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867
Db 1663 ACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1722
QY 868 AGGACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927
Db 1723 CGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1782
QY 928 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 987
Db 1783 CTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
QY 988 CTGAACCAAAATATTACAAATTCGCAAGACACCTCTCTGGAGGCGCTCTCTGAGAGGAC 1047
Db 1843 TTTCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1902
QY 1048 AGGACAAAGCGCT---CGGGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
Db 1903 CCAAGCAGAGGCTTGGCGGGGCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1962
QY 1105 TTTCTCTTAATTAAGTGGATGATCTCAATTAAGAAGATTTACTTCCCGCTTTTAAACCA 1164
Db 1963 TTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2022

QY 1165 ATGTGAGTGGGCCCCAACGAGTACGCGCACTTTGACCCGAGCTTTACCGAAGCCCTGTC 1224
 DB 2023 CAGTCCAGTCCGAGACTGACACCAAGGATTTTGTAGGAGGATTCACGGGCCAGATGATC 2082
 QY 1225 CCCAACTC 1232
 DB 2083 ACCATCAC 2090

RESULT 8

US-09-212-771-1

; Sequence 1, Application US/09212771

; Patent No. 5958773

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Lex M. Cowsett

; TITLE OF INVENTION: ANTISENSE MODULATION OF AKT-1 EXPRESSION

; FILE REFERENCE: PFS-0034

; CURRENT APPLICATION NUMBER: US/09/212,771

; CURRENT FILING DATE: 1998-12-16

; NUMBER OF SEQ ID NOS: 47

; SEQ ID NO 1

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (199)..(1641)

US-09-212-771-1

Query Match

Best Local Similarity 10.6%; Score 252; DB 2; Length 2610;

Matches 519; Conservative 0; Mismatches 380; Indels 9; Gaps 3;

QY 331 GACTTTCACCTCTTGAAGTGTGCGAAGGCGCAGTTTTCGAAAGGTTCTTCTACGAGA 390
 DB 643 GAGTTTGAGTACCTGAGCTGCTGGCAGAGGCACTTTCCGCAAGGTGATCTGTGTAAG 702
 QY 391 CACAAAGCAGAGAGTGTCTATGCGATGCAAGTTTTCAGAGAGAGCAATCTCTGAAA 450
 DB 703 GAGAGAGCCAGAGGCGGCGAGCTGTTCCACCTGTCGCGGAGAGCTGTCTCTCCGAG 762
 QY 451 AAGAAAGAGAGAGAGTATATGTCGAGCGGAGTGTCTGTTGCAAGATCTGAGAGC 510
 DB 763 AAGCAGAGGTGGCGCCACACACTCACGCA---GAACCGGCTCTCGAGAGCTCCAGGCAC 819
 QY 511 CCTTCTGCTGGGCGCTTCTCTTCTTCCAGACTGCTGACAAATTTGACTTCTCTCTA 570
 DB 820 CCTTCTGCTGAGGCGCTTCTCTTCTTCCAGAGCTGCTGAGAGAGAGTCTGCTGCTGATG 879
 QY 571 GACTACATATATGTTGAGAGTGTCTTCTTACATCTCCAGAGGAGAGGCTGCTCTTGAA 630
 DB 880 GAGTACGCAAGAGGCGGCGAGCTGTTCCACCTGTCGCGGAGAGCTGTCTCTCCGAG 939
 QY 631 CCACGGCTCTCTTCTATGCTGTGAATAGCCAGTCCCTTGGCTACCTGCTGATTC---A 687
 DB 940 GACCGGCGGCGCTTCTATGCGGCTGAGATTTGCTCAGCCCTGAGTACCTGCTGCTGCTG 999
 QY 688 CTGAACATCTGTTATGAGACTTAAACACAGAGATATTTTCTAGNTTTCACAGGAGAC 747
 DB 1000 AAGAAGCTGTGTACCGGAGCTCAAGCTGGAGAGCTGATGCTGACAGAGGAGCGGCAC 1059
 QY 748 ATTGCTCTTACTGATTTGCGACTCTGCAAGSAGAGACTTGAACACACAGCAGACATCC 807
 DB 1060 ATTAAGATTCAGAGACTTCTGGGCTGTGCAAGAGGAGGAGTCAAGAGCGGTGCCACCATG 1119
 QY 808 ACCTTCTGTCGCGCCGAGATFCTGCGACCTGAGGTGCTTCTAAGCAGGCTTATGAC 867
 DB 1120 ACCTTTTGGCGACACCTGAGTACCTGCGCCCGGAGGTGTGGAGGACATGATGATGCGC 1179
 QY 868 AGGACTGTGGAGTGTGCTGCGGAGCTGCTTCTATCAGATGCTGATGCGCTGCGC 927

DB 1180 CGTGCACTGGACTGCTGGGGCTGGCGCTGATGTATGTACGAGATGATGTGCTGCTGCTG 1239
 QY 928 CTTTCTTTATACCCGAACACAGCTGAAATGTACGACACATTTCTGAACAGCTCTCTGAG 987
 DB 1240 CCACTTCTACACACAGGACCATGAGAGCTTTTGTGGCTCATCTCTCATGAGAGAGATCCG 1299
 QY 988 CTGAACCAAAATATTACAAATTCGCAAGACACTCTCTGGAGGCGCTCTCTGCAAGAGAC 1047
 DB 1300 TTCCCGCGCAGCTTGTGTCGCGAGGCCAAGTCTTGTCTTTCAGGGCTGTCTCAAGAGAG 1359
 QY 1048 AGGACAAAGCGCT---CGGGGCCCAAGGATGACTTTCATGTAGATTAAGAGTCTATGCTTC 1104
 DB 1360 CCAAGCAGAGGCTTTGGCGGGCTTCCGAGAGCCGCCAAGAGATCTAGCAGCATCTCTTC 1419
 QY 1105 TTTCTCTTAATTAACCTGGGATGATCTCTATTAAAGAGATTTACTCCCTCTTTTAACTCA 1164
 DB 1420 TTTGCGGCTATGCTGTGGCAGCAGCTGTACGAGAGAGAGCTCAGCCCACTCTTCAAGAG 1479
 QY 1165 ATGTGAGTGGGCGCCCAAGCAGCTAGCGGCACCTTTCACCCCGAGTTTACCGAAGAGCTGTC 1224
 DB 1480 CAGGTCACTGCGAGACTGACACCAAGTATTTTGTATGAGGAGTTCACGGCCAGATGATC 1539
 QY 1225 CCCAACTC 1232
 DB 1540 ACCATCAC 1547

RESULT 9

US-09-091-058-1

; Sequence 1, Application US/09091058

; Patent No. 6054285

; GENERAL INFORMATION:

; APPLICANT: Hemmings, Brian A.

; APPLICANT: Frech, Matthias

; TITLE OF INVENTION: Screening Method

; FILE REFERENCE: 4-20683/A/20684/PCT

; CURRENT APPLICATION NUMBER: US/09/091,058

; CURRENT FILING DATE: 1998-06-10

; EARLIER APPLICATION NUMBER: PCT/EP96/04814

; EARLIER FILING DATE: 1996-11-05

; EARLIER APPLICATION NUMBER: 9525703.6

; EARLIER FILING DATE: 1995-12-15

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2610

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (199)..(1641)

US-09-091-058-1

Query Match

Best Local Similarity 10.6%; Score 252; DB 3; Length 2610;

Matches 519; Conservative 0; Mismatches 380; Indels 9; Gaps 3;

QY 331 GACTTTCACCTCTTGAAGTGTGCGAAGGCGCAGTTTTCGAAAGGTTCTTCTACGAGA 390
 DB 643 GAGTTTGAGTACCTGAGCTGCTGGCAGAGGCACTTTTCGCAAGGTGATCTGCTGCTGTAAG 702
 QY 391 CACAAGCAGAGAGAGTGTCTATGCAAGTCTTATGCAAGTCTTACAGAGAGAGCAATCTCTGAAA 450
 DB 703 GAGAGGCCACAGGCGCTACTACGCCATGAGATCTCTCAGAGAGGAGTATCTGTGCC 762
 QY 451 AAGAAAGAGAGAGAGCATATTATGTCGAGCGGAGATTTCTGTTGTAAGAGATGTGAAGACAC 510
 DB 763 AAGAGAGAGGTGGGCCCCACACACTCACCGA---GAACCGGCTGCTGACAGATCTCCAGCAC 819
 QY 511 CTTTCTGCTGCTGCGCTTCTCAGTCTTCTCCAGAGCTGCTGACAAATTTGACTTCTGCTCTA 570
 DB 820 CCTTCTCTACAGCCCTGAGTACTCTTTCAGAGCCCGCAGCAGCCGCTGCTGCTGCTGCTG 879

Db 1042 TTCAATTTCCATCGTCTGGGAAAGGGAGTTTGCAGAGGTGATGCTTGGCAGAGG 1101
QY 394 AAGGACAGAGTGTCTATGCACTCAAGTTTACAGAGAGCAATCCTGAGAAAG 453
Db 1102 AAGGACAGAGTGTCTATGCACTCAAGTTTACAGAGAGCAATCCTGAGAAAG 1161
QY 454 AAGGACAGAGTGTCTATGCACTCAAGTTTACAGAGAGCAATCCTGAGAAAG 513
Db 1162 GATGAGTGGAGTCAACATGTTAGAAAAGCGAGTCTTGGCCCTGCTTGACAAAACCCCG 1221
QY 514 TTCTGCTGGGCTTCACTCTCTTCCAGAGTCTGACAAATTTACTTGTCTAGAG 573
Db 1222 TTCTGCTGGGCTTCACTCTCTTCCAGAGTCTGACAAATTTACTTGTCTAGAG 1281
QY 574 TCAATTAATGTTGAGAGTGTCTTCCAGAGTCTGACAAATTTACTTGTCTAGAG 633
Db 1282 TATGTCACAGGTTGGGAGTCTTATGACCAATTCAGCAAGTAGGAAATTTANGAACCA 1341
QY 634 CGGGCTGCTTCTATGCTGCAATAGCCAGTCTTGGGCTACCTGCACTTCACTGAAC 693
Db 1342 CAAGCAGTATTTCTGCGGAGAGATTTCCATCGGATTTCTTCTTCATAAAGAGGA 1401
QY 694 ATCGTTTATAGAGTCTTAAACCCAGAGATTTTGTGATGATTCACAGGAGCAATTTGTC 753
Db 1402 ATCATTTATAGGATCTGAAGTTAGATAACGTCATGTTGGATTCAGAAAGGACATATCAA 1461
QY 754 CTACATGATTTGGGACTCTGCAAGGAGAGATTTGAACACACACAGCAACATCCACCTTC 813
Db 1462 ATTGCTGACTTTGGGATGTCAGAGGACATATGATGATGATGATGATGATGATGATGAT 1521
QY 814 TGTGGACCCCGAGTATCTCCACCTGAGGTGCTTCATAAGCAGCCTTATGACAGACT 873
Db 1522 TGTGGACTCCAGATATATATCCCGCAGAGATTAATCGCTTATCAGCCGCTATGAAATCT 1581
QY 874 GTGACTGTGTGCTGGGAGTCTGCTTGTATGATGATGCTGTATGCTGCTGCTGCTGCT 933
Db 1582 GTGACTGTGTGCTGGGAGTCTGCTTGTATGATGATGCTGTATGCTGCTGCTGCTGCT 1641
QY 934 TATAGCGGAACACAGTCAATGATGACAGACATTTGACAAAGCCTTCTCCAGCTGAAA 993
Db 1642 GATGTTGAGATGACAGAGGCTTATTCAGTCTATGAGGACACAGCTTCTCTATCCA 1701
QY 994 CCAATATTACAAATTCGCAAGACACCTCTGAGGAGGCTCTGCAAGAGGAGGACA 1053
Db 1702 AATCTCTGTCAGAGGCTGTTTCTATCTGCAAGGACTGATGACCAACACCCAGCC 1761
QY 1054 AAGCGCTCTGCGGAGGAGTCTATGAGGATGATGAGGATGATGATGATGATGATGATGAT 1110
Db 1762 AAGCGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1821
QY 1111 TTAATTAAGTGGATGATCTATTAATAAGAGATTAAGTGGGAGGAGGAGGAGGAGGAGGAG 1170
Db 1822 AGATCGACTGGGAAATCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1881
QY 1171 AGTGG 1175
Db 1882 TGTGG 1886

RESULT 13

US-09-772-647-3

; Sequence 3, Application US/09772647

; Patent No. 6521815

; GENERAL INFORMATION:

; APPLICANT: Verma, Ajit K

; APPLICANT: Reddig, Peter J

; APPLICANT: Jansen, Aaron P

; TITLE OF INVENTION: Animal Model System for Squamous Cell Carcinoma

; FILE REFERENCE: 960296.97613

; CURRENT APPLICATION NUMBER: US/09/772,647

; CURRENT FILING DATE: 2001-01-30

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2274

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: T7 tag and

; OTHER INFORMATION: mouse protein kinase C epsilon coding sequence

; NAME/KEY: CDS

; LOCATION: (61)..(2271)

; US-09-772-647-3

Query Match

Best Local Similarity 9.8%; Score 232.6; DB 4; Length 2274;

Matches 488; Conservative 0; Mismatches 394; Indels 9; caps 1;

QY 330 TGACTTTCACTTCTTGAAGTGTATGCACTCAAGTTTACAGAGAGCAATCCTGAA 449
Db 1278 TGAGTTCACTTCACTTGAAGTGTATGCACTCAAGTTTACAGAGAGCAATCCTGAA 1337
QY 390 ACACAGAGCAGAGAGTGTCTATGCACTCAAGTTTACAGAGAGCAATCCTGAA 449
Db 1338 ACTCAAGGCAAGAGTGTATGCACTCAAGTTTACAGAGAGCAATCCTGAA 1497
QY 450 AAGAGAGAGGAGAGATATATGTCGGAGCGGAATGTTCTGTGAGAGATCTGAAAGA 509
Db 1398 AGACGATGATGTGGACTCCACATGACAGAGAGAGGATTTTGGCTCTGGCTCGGAACA 1457
QY 510 CCTTTTCCTGTTGGGCTTCACTTCTTTCAGACTGCTGACAAATTCGACTTGGACT 569
Db 1458 CCTTATCTAACCCCACTTATGCTGCTTCCAGACCAAGAGCCGCTTCTTCTGCAAT 1517
QY 570 AGACTACATTAATGTTGGAGAGTGTCTTACCATCTCCAGAGGAGAGCTGTTCTGGA 629
Db 1518 GGAATATGTAATGTTGGAGAGTGTCTTCCAGATTCAGCGGTCGCCGAAATTTGATGA 1577
QY 630 ACCAGCGGCTGCTTCTTATGCTGCTGATAGCAGTCCCTTGGGCTACGAGTCACT 689
Db 1578 GCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1637
QY 690 GAACATGCTTTATAGAGACTTAAACACAGAGATATTTTGTAGATTCATAGGATACAT 749
Db 1638 TGGAGTATCTACAGAGATTTGAATCTGCAACATCTCTTCTAGATCAGAGAGCTG 1697
QY 750 TGTCTTCTAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
Db 1698 CAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1757
QY 810 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
Db 1758 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1817
QY 870 GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929
Db 1818 CTCAGTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1877
QY 930 TTTTATAGCCGAAACACAGCTGAAATGTACCAACATCTTGAACAGCTCTTCCAGCT 989
Db 1878 CTTTGAAGCTGACAGAGAGAGTCTTCTTGAATCCATCTTCTATGATGATGATGATGAT 1937
QY 990 GAAACCAATATTACAAATTCGCGAGAGACCTCTCTGAGGAGGCTCTCTGAGAGAGAG 1049
Db 1938 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1997
QY 1050 GACAAAGCGCTCGG-----GGCCCAAGATGACTTCTATGAGATTAAGAGCTATGT 1100
Db 1998 GCACAGCGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2057
QY 1101 CT 1160
Db 2058 ATTCTTCAAGAGGATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2117
QY 1161 CCCAATGTGAGTGGGCTTCAAGAGCTATCGGCACTTTTGACCCCGAGTTTAC 1211

Search completed: August 11, 2003, 06:10:59
Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 11, 2003, 08:36:05 ; Search time 514.182 Seconds
(without alignments)
756.292 Million cell updates/sec

Title: US-10-000-039A-4
Perfect score: 85
Sequence: 1 EAFGLFVAPPTDSFL 16

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delcp 6.0, Delcxt 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-O=/cpn2-1/USPTO.spool/US10000039/runat_01082003_085419_27521/app.query.fasta_1.661
-DB=EST -FPM=fastap -SUFFIX=rst -MINMATCH=0.1 -LOECL=0 -LOEPEXT=0
-UNIT=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10000039 -CCN_1_1.6635 -runat_01082003_085419_27521 -NCPU=6 -ICPU=3
-NO_MMAB -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_nam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	85	100.0	344	10	BE001610
2	85	100.0	346	9	AA377138
3	85	100.0	376	12	B1032559
4	85	100.0	404	14	H72913
5	85	100.0	420	14	R86234
6	85	100.0	427	10	BE814167
7	85	100.0	448	9	A1188198
8	85	100.0	458	14	CB114428
9	85	100.0	460	14	R35214
10	85	100.0	465	10	BE002651
11	85	100.0	465	12	BE004999
12	85	100.0	487	10	BF950844
13	85	100.0	505	12	BM707970
14	85	100.0	530	12	BM756001
15	85	100.0	544	13	BQ083706
16	85	100.0	552	14	CB216289
17	85	100.0	557	12	BM705333
18	85	100.0	572	12	BM753618
19	85	100.0	582	9	AV702971
20	85	100.0	586	12	B1003243
21	85	100.0	608	14	CB216184
22	85	100.0	620	10	BF350342
23	85	100.0	630	14	CB216592
24	85	100.0	641	9	AW378705
25	85	100.0	654	9	AL040857
26	85	100.0	659	10	BE178455
27	85	100.0	661	10	BE178442
28	85	100.0	663	10	BE178469
29	85	100.0	671	12	BG770190
30	85	100.0	703	10	BG485236
31	85	100.0	721	12	BG778994
32	85	100.0	736	10	BG575325
33	85	100.0	765	12	BG763506
34	85	100.0	767	14	CH000714
35	85	100.0	791	10	BF691391
36	85	100.0	826	14	CD517205
37	85	100.0	834	12	B1335545
38	85	100.0	847	13	BQ212589
39	85	100.0	851	12	B1764996
40	85	100.0	859	13	BQ680126
41	85	100.0	873	10	BG756781
42	85	100.0	876	10	BG122546
43	85	100.0	885	10	BG753783
44	85	100.0	898	13	BQ687393
45	85	100.0	899	12	BG769533

ALIGNMENTS

RESULT 1
BE001610
LOCUS PM2-BN0080-270400-003-a12 BN0080 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE001610
ACCESSION BE001610.1 GI:8261843
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 344)
AUTHORS Dias Neto,E., Verjovsky-Almeida,S., Briones,M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-pm2-BN0080-270
400-003-ai2&t3-2000-04-27&t4-1)

Seq primer: puc 18 forward

High quality sequence start: 30

High quality sequence stop: 344

Location/Qualifiers

1..344

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="BN0080"

/dev_stage="Adult"

/note="Organ: breast_normal; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 75 a 92 c 76 g 101 t

ORIGIN

Alignment Scores:

Pred. No.: 0.003 Length: 344

Score: 85.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 10 Gaps: 0

US-10-000-039A-4 (1-16) x BE001610 (1-344)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProPThrAspSerPheLeu 16

Db 227 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGAGCTCTTCTC 274

RESULT 2

AA377138

LOCUS

DEFINITION

AA377138

ACCESSION

AA377138.1 GI:2029466

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 346)

ADAMS, M.D., KERLAVAGE, A.R., FLEISCHMANN, R.D., FULDNER, R.A., BULT,

C.J., LEE, N.H., KIRKNESS, E.F., WEINSTOCK, K.G., GOCAYNE, J.D., WHITE,

O., SUTTON, G., BLAKE, J.A., BRANDON, R.C., MAN-RAI, C., CLAYTON, R.A.,

CLINE, T.R., COTTON, M.D., EARLE-HUGHES, J., FINE, L.D., FITZGERALD

EST89672 Small intestine I Homo sapiens cDNA 5' end similar to

similar to serine/threonine kinase, mRNA sequence.

AA377138

ACCESSION

AA377138.1 GI:2029466

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 346)

ADAMS, M.D., KERLAVAGE, A.R., FLEISCHMANN, R.D., FULDNER, R.A., BULT,

C.J., LEE, N.H., KIRKNESS, E.F., WEINSTOCK, K.G., GOCAYNE, J.D., WHITE,

O., SUTTON, G., BLAKE, J.A., BRANDON, R.C., MAN-RAI, C., CLAYTON, R.A.,

L.M., Fitzhugh, W.M., Fritschman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Morenno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelli, L., S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Raymond, L., Kunsch, C., Hung, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillman, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Other ESTs: THC173211

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1..346

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="adult"

/clone_lib="Small intestine I"

/note="Organ: small intestine; Vector: pBluescript SK-"

BASE COUNT 80 a 93 c 74 g 97 t

ORIGIN

Alignment Scores:

Pred. No.: 0.00301 Length: 346

Score: 85.00 Matches: 16

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-000-039A-4 (1-16) x AA377138 (1-346)

Qy 1 GluAlaPheLeuGlyPheSerTyrAlaProPThrAspSerPheLeu 16

Db 157 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGAGCTCTTCTC 204

RESULT 3

AA377138

LOCUS

DEFINITION

AA377138

ACCESSION

AA377138.1 GI:14439185

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 376)

DIAS, N.E.O., E., GARCIA CORREA, R., VERJOVSKI-ALMEIDA, S., BRIONES, M.R.,

NAGAI, M.A., DA SILVA, W. JR., ZAGO, M.A., BORDIN, S., COSTA, F.F.,

GOIDMAN, G.H., CARVALHO, A.F., MATSUKUMA, A., BATA, G.S., SIMPSON, D.H.,

BRUNSTEIN, A., DEOLIVEIRA, P.S., BUCHER, P., JONGENEEL, C.V., O'HARE

EST89672 Small intestine I Homo sapiens cDNA, mRNA sequence.

AA377138

ACCESSION

AA377138.1 GI:14439185

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILML at:
 www.bio.lnli.gov/bbrp/image/image.html
 Insert length: 284 Std Error: 0.00
 Seq primer: -400P from Gibco.

FEATURES

source

1. .448

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:1734465"

/sex="male"

/lab_host="DH10B"

/clone_lib="Soares_testis_NHT"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTGAAGTGGAGCGGCCCAATTTTCTTTTCTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cots, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 143 a 105 c 93 g 107 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.00391 Length: 448
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-000-039a-4 (1-16) x A1188198 (1-448)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16

|||||

Db 448 GAGGCTTCTAGGCTTTTCTATCGCGCTCCACGGACTCTTTCCTC 401

RESULT 8

CB114428

LOCUS

DEFINITION

K-EST0158063 L5HLK1s1 Homo sapiens cDNA clone L5HLK1s1-1-A07 5',

mRNA sequence.

CB114428

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 458)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsungemall.kr@kribb.re.kr

Plate: 1 row: A column: 07

High quality sequence stop: 458.

Location/Qualifiers

1. .458

/organism="Homo sapiens"

FEATURES

source

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="L5HLK1s1-1-A07"

/sex="M"

/cell_line="HLK-1"

/lab_host="Top10F"

/clone_lib="L5HLK1s1"

/note="Organ: Liver; Vector: pCNS-D2; Site.1: EcoRI;

Site.2: NotI; The poly (A)+ RNA was dephosphorylated with

bacterial alkaline phosphatase (BAP) and then dephosphorylated

with tobacco acid pyrophosphatase (TAP). The dephosphorylated

intact mRNA was ligated with DNA-RNA linker including

EcoRI site by treatment of T4 RNA ligase and the first

strand cDNA was synthesized from oligo dT-selected mRNA by

priming with dT-tailed vector. The dT-tailed vector was

adjusted to have about 60nt. The cDNA vector was

circularized with E. coli DNA ligase after digestion of

EcoRI which site is also included in vector. An RNA strand

converted to a DNA strand by Okayama-Berg method. The

obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F' by electroporation method.

The cDNA libraries constructed by this method are

full-length enriched cDNA library. After analyzing and

sequencing about 2,000 - 3,000 colonies in original cDNA

library, the abundant cDNAs were selected and amplified by

PCR reaction using vector region primer including T7

promotor as 5' primer and N(dT)14 as 3' primer. The PCR

products were used as template for synthesis of

biotinylated single stranded RNA by in vitro transcription

reaction. The synthesized RNA probes were hybridized with

antisense single stranded cDNAs prepared from original

library and incubated with avidin-gel. After removing

DNA-RNA hybrids by centrifuge, the subtracted cDNA

libraries were constructed by transformation of the

remaining DNA into competent cells E. coli Top10F' with

electroporation method."

BASE COUNT 117 a 128 c 98 g 115 t

ORIGIN

Alignment Scores:
 Pred. No.: 0.00399 Length: 458
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-10-000-039a-4 (1-16) x CB114428 (1-458)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProProThrAspSerPheLeu 16

|||||

Db 357 GAGGCTTCTAGGCTTTTCTATCGCGCTCCACGGACTCTTTCCTC 404

RESULT 9

R35214

LOCUS

DEFINITION

IMAGE:136701 5' similar to SP:A48094 A48094 SERUM AND

GLUCOCORTICOID-REGULATED KINASE, SGK ;, mRNA sequence.

R35214

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 460)

Hallier,L., Clark,N., Dubucq,T., Elliston,K., Hawkins,M., Holman

,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,

Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston

,R., Williamson,A., Wohldmann,P. and Wilson,R.

The WashU-Merck EST Project

Unpublished

R35214 460 bp mRNA linear EST 02-MAY-1995

Yh87b11.r1 Soares placenta Nb2HP Homo sapiens cDNA clone

IMAGE:136701 5' similar to SP:A48094 A48094 SERUM AND

GLUCOCORTICOID-REGULATED KINASE, SGK ;, mRNA sequence.

R35214

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 460)

Hallier,L., Clark,N., Dubucq,T., Elliston,K., Hawkins,M., Holman

,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,

Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston

,R., Williamson,A., Wohldmann,P. and Wilson,R.

The WashU-Merck EST Project

Unpublished

tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

source
1. 465
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DW0-agg-g-13-0-UI"
/tissue_type="lens"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DW0"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-DW0 is a cDNA library containing the following tissue(s): lens. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CATTAGGGA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
BASE COUNT 101 a 124 c 101 g 139 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00405 Length: 465
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-10-000-039a-4 (1-16) x BM704999 (1-465)

Qy 1 GluAlaPheLeuGlyPheSerTyraAlaProThrAspSerPheLeu 16
|||||
Db 157 GAGGCTTCTTAGGCTTTCTATGCGCTCCACGAGACTCTTCCCTC 204

RESULT 12
BF950844 487 bp mRNA linear EST 22-JAN-2001
LOCUS
DEFINITION RC3-NN1182-061100-022-a03 NN1182 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF950844
VERSION BF950844.1 GI:12368119
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 487)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC3&l2=RC3-NN1182-061100-022-a03&t3=2000-11-06&l4=1)
Seq primer: puc 18 forward
High quality sequence stop: 487.
Location/Qualifiers
1. 487
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN1182"
/note="Organ: nervous normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 106 a 118 c 119 g 144 t
ORIGIN

FEATURES

source
1. 487
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN1182"
/note="Organ: nervous normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 106 a 118 c 119 g 144 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00425 Length: 487
Score: 85.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-10-000-039a-4 (1-16) x BF950844 (1-487)

Qy 1 GluAlaPheLeuGlyPheSerTyraAlaProThrAspSerPheLeu 16
|||||
Db 76 GAGGCTTCTTAGGCTTTCTATGCGCTCCACGAGACTCTTCCCTC 123

RESULT 13
BM707970 505 bp mRNA linear EST 28-FEB-2002
LOCUS
DEFINITION UI-E-C11-aft-k-17-0-UI.r1 UI-E-C11 Homo sapiens cDNA clone
UI-E-C11-aft-k-17-0-UI 5', mRNA sequence.

ACCESSION BM707970
VERSION BM707970.1 GI:19021228
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 505)
Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB

Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEHRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

```
1..505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-Cil-aft-k-17-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="GH108 (life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-Cil"
/note="organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-Cil is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(qT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NE1)."
BASE COUNT      124 a 143 c 117 g 121 t
ORIGIN
```

Alignment Scores:
 Pred. No.: 0.0044 Length: 505
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-10-000-039A-4 (1-16) x BM707970 (1-505)

```
OY 1 GUAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 16
|||||
Db 455 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGAGACTCTTTCCTC 502
```

```
RESULT 14
BM756001      530 bp mRNA linear EST 04-MAR-2002
LOCUS      K-EST0034163 SISNU5 Homo sapiens cDNA clone S1SNUS-38-E04 5', mRNA
DEFINITION      sequence.
ACCESSION      BM756001
VERSION      BM756001.1 GI:19085616
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 530)
AUTHORS      Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL      Unpublished
COMMENT      Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 38 row: E column: 04
```

FEATURES

High quality sequence stop: 530.
 Location/Qualifiers

```
1..530
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S1SNUS-38-E04"
/sex="F"
/tissue_type="Asclites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-5"
/lab_host="Top10F"
/clone_lib="S1SNUS"
/note="organ: Stomach; Vector: pcNS; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then dephapped
with tabacco acid pyrophosphatase (TAP). The deapped
intact mRNA was ligated with DNA-RNA linker including EcoK
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
BASE COUNT      125 a 140 c 120 g 145 t
ORIGIN
```

Alignment Scores:
 Pred. No.: 0.00462 Length: 530
 Score: 85.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-10-000-039A-4 (1-16) x BM756001 (1-530)

```
OY 1 GUAlaPheLeuGlyPheSerTyraAlaProProThrAspSerPheLeu 16
|||||
Db 290 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGAGACTCTTTCCTC 337
```

```
RESULT 15
BQ083706      544 bp mRNA linear EST 04-APR-2002
LOCUS      K-EST0146418 S14K402 Homo sapiens cDNA clone S14K402-63-H01 5',
DEFINITION      mRNA sequence.
ACCESSION      BQ083706
VERSION      BQ083706.1 GI:19942792
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 544)
AUTHORS      Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL      Unpublished
COMMENT      Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: H column: 01
High quality sequence stop: 544.
```

FEATURES
source

Location/Qualifiers
 1. .544
 /organism="Homo Sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S14K402-63-R01"
 /cell_line="K402"
 /lab_host="Top10F"
 /clone_lib="S14K402"
 /note="Organ: Stomach; Vector: pTZ18Rp1; Site_1: EcoRI;
 Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

BASE COUNT 116 a 140 C 129 g 159 t
 ORIGIN

Alignment Scores:

Pred. No.:	0.00475	Length:	544
Score:	85.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0

US-10-000-039A-4 (1-16) x BQ083706 (1-544)

QY 1 GluAlaPheLeuGlyPheSerTyrAlaProGthrAspSerPheLeu 16
 |||
 Db 136 GAGGCTTCTAGGCTTTTCCTATGCGCTCCACGGACTCTTTCCTC 183

Search completed: August 11, 2003, 11:27:48
 Job time : 515.182 secs